D.A. RAIKOV

ANALYSE MATHÉMATIQUE MULTIDIMENSIONNELLE

D. RAÏKOV

ANALYSE MATHÉMATIQUE MULTIDIMENSIONNELLE



ÉDITIONS MIR MOSCOU

TABLE DES MATIÈRES

Première partie

INTRODUCTION À L'ANALYSE MULTIDIMENSIONNELLE

CHAPITRE PREMIER

ESPACES MÉTRIQUES ET TOPOLOGIQUES

§ 1. Espace euclidien de dimension n

L'analyse multidimensionnelle a pour objet d'étude les espaces euclidiens de dimension finie quelconque. Dans ce paragraphe nous introduisons la notion d'*espace euclidien de dimension n* et les notions de *produit scalaire*, de *norme* et de *métrique* sur lesquelles elle est fondée.

1. Espace vectoriel $\mathbf{R}^n$

Le *produit d'une famille de n ensembles* est l'ensemble de tous les n-uples possibles d'éléments pris dans chaque ensemble de la famille. On désigne le produit d'une famille de n ensembles $E_1, \ldots, E_n$ par $E_1 \times \ldots \times E_n$. Si $E_1 = \ldots = E_n = E$ (i.e. la famille se compose de n exemplaires de l'ensemble E), on remplace la notation $\underbrace{E \times \ldots \times E}_{n \text{ fois}}$ par E^n. *Le produit $E_1 \times \ldots \times E_n$ d'ensembles non vides $E_1, \ldots, E_n$ n'est pas vide.* En effet, si $n = 1$, ce produit coïncide avec E_1 et n'est pas vide par hypothèse. Si l'on a déjà établi que le produit de n ensembles $E_1, \ldots, E_n$ est non vide, i.e. il existe $(x_1, \ldots, x_n) \in E_1 \times \ldots \times E_n$, et si l'ensemble E_{n+1} n'est pas vide, i.e. il existe $x_{n+1} \in E_{n+1}$, alors $(x_1, \ldots, x_n, x_{n+1}) \in E_1 \times \ldots \times E_n \times E_{n+1}$, de sorte que le produit des ensembles $E_1, \ldots, E_n, E_{n+1}$ n'est pas vide non plus *). Les éléments du produit $E_1 \times \ldots \times E_n$ s'écrivent sous forme de n-uples $(x_1, \ldots, x_n)$, avec $x_k \in E_k$ $(k = 1, \ldots, n)$. Donc, la place de l'élément dans un n-uple correspond à l'ensemble dont il est choisi.

En particulier, $\mathbf{R}^n$ *est l'ensemble de tous les n-uples de nombres réels.* Ses éléments $(x_1, x_2, \ldots, x_n)$ sont habituellement appelés *points*, et les nombres $x_1, x_2, \ldots, x_n$, première, deuxième, ..., n-ième *coordonnées* du point $(x_1, x_2, \ldots, x_n)$. Les points seront souvent désignés par une seule lettre : par une lettre minuscule $x = (x_1, \ldots, x_n)$ dans les expressions

*) Le problème d'un produit non vide d'une famille infinie d'ensembles non vides est assez délicat.

analytiques ou par une majuscule, par exemple M suivie parfois de coordonnées, $M(x_1, \ldots, x_n)$, dans les expressions géométriques. Pour $n = 1$, 2, 3 et, parfois, pour $n = 4$, la numérotation est habituellement omise : pour $\mathbf{R}^1 = \mathbf{R}$ on utilise une seule lettre, pour $\mathbf{R}^2$ le point est désigné par (x, y), pour $\mathbf{R}^3$ par (x, y, z) et pour $\mathbf{R}^4$ (parfois) par (x, y, z, t). Le symbole $\mathbf{R}^2$ représentera le plan de coordonnées, et $\mathbf{R}^3$, l'espace des coordonnées.

Il est à noter que $\mathbf{R}^n$ n'est pas seulement un ensemble : il possède encore certaines structures mathématiques standard. Tout d'abord on peut le munir des opérations d'*addition coordonnée à coordonnée* et de *multiplication par les nombres réels* (*scalaires*). On pose par définition

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

ce qui fait de $\mathbf{R}^n$ un *espace vectoriel sur le corps* $\mathbf{R}$ (*espace vectoriel réel*). L'élément nul de $\mathbf{R}^n$ est le point $0 = (0, \ldots, 0)$. Pour $n = 2$ et $n = 3$, ces opérations correspondent à l'addition des rayons vecteurs des points correspondants du plan ou de l'espace et à leur multiplication par les nombres réels. Pour cette raison, les éléments des ensembles $\mathbf{R}^2$ et $\mathbf{R}^3$ peuvent être aussi identifiés aux vecteurs. (D'où le terme « espace vectoriel »). Par souci de commodité, nous appellerons les éléments de l'espace $\mathbf{R}^n$ tantôt points, tantôt vecteurs.

Il existe dans l'espace vectoriel $\mathbf{R}^n$ une *base standard* formée par les vecteurs

$$e^1 = (1, 0, \ldots, 0), \quad e^2 = (0, 1, \ldots, 0), \ldots, e^n = (0, \ldots, 0, 1) ;$$

leur notation générale est $e^k = (\delta_1^k, \delta_2^k, \ldots, \delta_n^k)$, où δ_i^k est le *symbole de Kronecker*, i.e.

$$\delta_i^k = \begin{cases} 1 & \text{si } i = k, \\ 0 & \text{si } i \neq k. \end{cases}$$

En effet, $(x_1, \ldots, x_n) = x_1 e^1 + \ldots + x_n e^n$ et cette représentation du point $(x_1, \ldots, x_n)$ sous forme d'une combinaison linéaire des vecteurs $e^1, \ldots, e^n$ est unique. Donc, *$\mathbf{R}^n$ est un espace vectoriel de dimension n sur* $\mathbf{R}$.

2. Produit scalaire

Il existe aussi dans $\mathbf{R}^n$ un produit scalaire standard.

DÉFINITION 1. Le *produit scalaire* sur un espace vectoriel réel X est une fonction qui associe à chaque couple (u, v) de vecteurs de X un nombre réel (nous le désignerons par $\langle u, v \rangle$) de façon que soient satisfaites les conditions suivantes (*axiomes du produit scalaire*) :

1) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ pour tout $(x, y, z) \in X^3$;
2) $\langle \lambda x, z \rangle = \lambda \langle x, z \rangle$ pour tout $(x, z, \lambda) \in X \times X \times \mathbf{R}$;
3) $\langle x, y \rangle = \langle y, x \rangle$ pour tout $(x, y) \in X^2$;
4) $\langle x, x \rangle > 0$ si $x \in X$ et $x \neq 0$.

Par récurrence on étend la première condition à toutes les sommes finies :

$$\langle x^1 + \ldots + x^k, z \rangle = \langle x^1, z \rangle + \ldots + \langle x^k, z \rangle$$

pour tout $k \in \mathbf{N}$ et tout $(x^1, \ldots, x^k, z) \in X^{k+1}$. La *linéarité* du produit scalaire suivant le premier facteur, quelle que soit la valeur fixée du deuxième, découle de la première et de la deuxième condition :

$$\langle \lambda_1 x^1 + \ldots + \lambda_k x^k, y \rangle = \lambda_1 \langle x^1, y \rangle + \ldots + \lambda_k \langle x^k, y \rangle$$

pour tous les $k \in \mathbf{N}$, $(x^1, \ldots, x^k, y) \in X^{k+1}$ et tout $(\lambda_1, \ldots, \lambda_k) \in \mathbf{R}^k$. Jointe à la troisième condition, cette propriété entraîne la linéarité du produit scalaire suivant le deuxième facteur pour chaque valeur fixée du premier. Donc, $\langle u, v \rangle$ est une *fonction bilinéaire.*

Pour $\lambda = 0$, la deuxième condition implique $\langle 0, z \rangle = 0$ pour tout $z \in X$, ce qui signifie en vertu de la troisième condition que $\langle z, 0 \rangle = 0$ pour tout $z \in X$. En particulier, $\langle 0, 0 \rangle = 0$. Par conséquent, la quatrième condition signifie que $\langle x, x \rangle \geqslant 0$ pour tout $x \in X$, et $\langle x, x \rangle = 0 \Leftrightarrow$ $\Leftrightarrow x = 0$.

On voit facilement que la formule

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k \tag{1}$$

définit le produit scalaire des vecteurs $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots \ldots, y_n)$ dans $\mathbf{R}^n$. Pour $n = 1$, c'est un produit ordinaire ; $\langle x, y \rangle = xy$. Pour $n = 2$ et $n = 3$, le produit scalaire de deux vecteurs (défini géométriquement comme produit des modules de ces vecteurs par le cosinus de l'angle qu'ils font) est calculé en coordonnées cartésiennes rectangulaires d'après cette formule, c'est-à-dire comme somme des produits de coordonnées correspondantes de ces vecteurs. Dans le cas général, la formule (1) définit le *produit scalaire standard* sur $\mathbf{R}^n$. L'espace $\mathbf{R}^n$ possède une infinité d'autres produits scalaires, par exemple $\sum_{k=1}^{n} \varrho_k x_k y_k$, où ϱ_k sont des coefficients positifs fixés quelconques.

3. Inégalité de Cauchy-Bouniakovski

THÉORÈME 1. *Le produit scalaire $\langle x, y \rangle$ sur un espace vectoriel réel X satisfait à l'inégalité*

$$\langle x, y \rangle^2 \leqslant \langle x, x \rangle \langle y, y \rangle. \tag{2}$$

Démonstration. En vertu des axiomes du produit scalaire, on a pour tout $\lambda \in \mathbf{R}$

(3) $$\lambda^2\langle x, x\rangle - 2\lambda\langle x, y\rangle + \langle y, y\rangle = \langle \lambda x - y, \lambda x - y\rangle \geqslant 0.$$

Si $\langle x, x\rangle \neq 0$, alors pour $\lambda = \dfrac{\langle x, y\rangle}{\langle x, x\rangle}$ cela donne

$$-\frac{\langle x, y\rangle^2}{\langle x, x\rangle} + \langle y, y\rangle \geqslant 0,$$

d'où découle (2). Si, par contre, $\langle x, x\rangle = 0$, alors (3) se met pour tout $\lambda \in \mathbf{R}$ sous la forme

(4) $$-2\lambda\langle x, y\rangle + \langle y, y\rangle \geqslant 0,$$

ce qui n'est possible que si $\langle x, y\rangle = 0$, de sorte que l'inégalité (2) est vérifiée pour ce cas aussi (avec le signe d'égalité).

REMARQUE. *Le signe d'égalité dans* (2) *a lieu si et seulement si les vecteurs x, y sont linéairement dépendants.* En effet, soient x et y des vecteurs linéairement dépendants. Si $x = 0$, l'égalité dans (2) est évidente. Si $x \neq 0$ et, par suite de la quatrième condition, $\langle x, x\rangle \neq 0$, alors $y = \lambda x$, de sorte que $\langle x, y\rangle = \lambda\langle x, x\rangle$, d'où $\lambda = \dfrac{\langle x, y\rangle}{\langle x, x\rangle}$, et

$$\langle y, y\rangle = \lambda\langle x, y\rangle = \frac{\langle x, y\rangle^2}{\langle x, x\rangle}$$

entraîne $\langle x, y\rangle^2 = \langle x, x\rangle\langle y, y\rangle$. Inversement : si la dernière égalité est juste et si $x \neq 0$, i.e. $\langle x, x\rangle \neq 0$, alors (3) devient égalité pour $\lambda = \dfrac{\langle x, y\rangle}{\langle x, x\rangle}$, i.e. $\langle \lambda x - y, \lambda x - y\rangle = 0$ ou, ce qui revient au même, $y = \lambda x$. La dépendance linéaire de x et y pour $x = 0$ est évidente.

La relation (2) s'appelle *inégalité de Cauchy-Bouniakovski.* Pour le produit scalaire standard (1) dans $\mathbf{R}^n$ elle prend la forme suivante

(5) $$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leqslant \sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} y_k^2.$$

L'inégalité (5) a été obtenue pour la première fois par Cauchy. Citons encore une interprétation de l'inégalité (2). Soit X l'ensemble de toutes les fonctions réelles continues sur un intervalle donné non dégénéré $[a, b]$; X est un espace vectoriel réel pour les opérations d'addition des fonctions et de multiplication d'une fonction par un nombre. Pour chaque couple (x, y) de fonctions de X posons

$$\langle x, y\rangle = \int_a^b x(t)\, y(t)\, \mathrm{d}t.$$

On voit aisément que c'est un produit scalaire dans X (la quatrième condition découle de la continuité des fonctions qui forment X : si $x(t) \not\equiv 0$, alors $\int_a^b x^2(t)\,dt > 0$ en vertu de la continuité de x. L'inégalité (2) prend la forme

$$\left(\int_a^b x(t)\,y(t)\,dt\right)^2 \leqslant \int_a^b x^2(t)\,dt \int_a^b y^2(t)\,dt.$$

Cette inégalité a été obtenue pour la première fois par Bouniakovski.

4. Norme

Le produit scalaire sur un espace vectoriel réel est naturellement lié à la notion de norme.

DÉFINITION 2. La *norme* sur un espace vectoriel réel X est une fonction qui associe à chaque point $x \in X$ le nombre réel $\|x\|$ (on lit : norme de x) de façon que soient satisfaites toutes les conditions suivantes (*axiomes de la norme*) :

1) $\|x + y\| \leqslant \|x\| + \|y\|$ pour tout $(x, y) \in X^2$;
2) $\|\lambda x\| = |\lambda|\,\|x\|$ pour tout $(x, \lambda) \in X \times \mathbf{R}$;
3) $\|x\| \neq 0$ si $x \in X$ et $x \neq 0$.

Il découle de l'axiome 2 que $\|0\| = 0$. En effet, $\|0\| = \|0 \cdot 0\| = |0|\,\|0\| = 0$. Nous obtenons alors des axiomes 1 et 2

$$0 = \|x + (-x)\| \leqslant \|x\| + \|x\| = 2\|x\|,$$

d'où $\|x\| \geqslant 0$ pour tout $x \in X$. Associée à l'axiome 3, l'égalité $\|0\| = 0$ entraîne $\|x\| = 0 \Leftrightarrow x = 0$.

THÉORÈME 2. *Soit X un espace vectoriel réel muni du produit scalaire $\langle x, y\rangle$. Alors*

$$\sqrt{\langle x, x\rangle} \qquad (6)$$

est une norme sur X.

Démonstration. L'expression (6) a du sens puisque $\langle x, x\rangle \geqslant 0$. Vérifions les axiomes de la norme.

1. Les axiomes 1 et 3 du produit scalaire et l'inégalité de Cauchy-Bouniakovski entraînent

$$\langle x + y, x + y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leqslant$$
$$\leqslant \langle x, x\rangle + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + \langle y, y\rangle =$$
$$= \left(\sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}\right)^2,$$

de sorte que $\sqrt{\langle x + y, x + y\rangle} \leqslant \sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}$.

2. En vertu des axiomes 2 et 3 du produit scalaire,

$$\sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda^2\langle x, x\rangle} = |\lambda|\,\sqrt{\langle x, x\rangle}.$$

3. Si $x \in X$ et $x \neq 0$, alors $\langle x, x \rangle > 0$ en vertu de l'axiome 4 du produit scalaire, si bien que $\sqrt{\langle x, x \rangle} \neq 0$.

Nous dirons que (6) est une *norme engendrée par le produit scalaire.*

EXERCICE 1. Montrer que $\sqrt{\langle x + y, x + y \rangle} = \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}$ si et seulement si $y = \alpha x$, avec $\alpha \geqslant 0$, ou si $x = 0$.

EXERCICE 2. Montrer que la norme engendrée par le produit scalaire vérifie l'identité

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{7}$$

REMARQUE. On peut montrer que la réciproque est aussi juste : la norme qui vérifie l'identité (7) est engendrée par un produit scalaire défini par $\frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$ *).

Pour le produit scalaire standard sur $\mathbf{R}^n$ la norme (6) prend la forme

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}. \tag{8}$$

Nous l'appellerons *norme euclidienne sur* $\mathbf{R}^n$.

5. Métrique

La norme sur un espace vectoriel réel engendre naturellement la distance.

DÉFINITION 3. La *métrique* (ou la *distance*) sur un ensemble non vide X est une fonction qui associe à chaque couple (x, y) d'éléments de X un nombre réel (noté $d(x, y)$ et appelé *distance de x à y*) de façon que soient satisfaites les conditions suivantes (*axiomes de la métrique*) :

1) $d(x, y) \leqslant d(x, z) + d(y, z)$ pour tout $(x, y, z) \in X^3$.

2) $d(x, y) = 0 \Leftrightarrow x = y$.

En posant $z = x$ dans l'inégalité de l'axiome 1 et en prenant en considération l'axiome 2, nous obtenons $d(x, y) \leqslant d(x, x) + d(y, x) = d(y, x)$; pour les mêmes raisons $d(y, x) \leqslant d(x, y)$. Par conséquent,

$$d(x, y) = d(y, x) \text{ pour tout } (x, y) \in X^2$$

(la distance de x à y est égale à celle de y à x). D'autre part, en posant $y = x$ dans l'inégalité de l'axiome 1 et en tenant compte de l'axiome 2, nous obtenons $0 = d(x, x) \leqslant 2d(x, z)$, d'où

$$d(x, z) \geqslant 0 \text{ pour tout } (x, z) \in X^2.$$

L'inégalité de l'axiome 1 ou, ce qui maintenant revient au même, l'inégalité

$$d(x, y) \leqslant d(x, z) + d(z, y) \tag{9}$$

est appelée *inégalité triangulaire* (« la longueur d'un côté du triangle ne dépasse pas la somme de ses deux autres côtés »).

---

*) Voir KOLMOGOROV A., FOMINE S., *Eléments de la théorie des fonctions et de l'analyse fonctionnelle*, Moscou, éd. Mir, 1977 (chap. III, § 4, n° 8).

THÉORÈME 3. *Soit* $\|x\|$ *une norme sur un espace vectoriel réel X. Alors*

$$d(x, y) = \|x - y\| \tag{10}$$

est une métrique sur X.

Démonstration. 1) Vu que $x - y = (x - z) + (z - y)$, on a

$$\|x - y\| \leqslant \|x - z\| + \|z - y\| = \|x - z\| + \|y - z\|.$$

2) $\|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

En particulier, la norme euclidienne sur $\mathbf{R}^n$ engendre la *métrique euclidienne*

$$d(x, y) = \sqrt{\sum_{k=1}^{n}(x_k - y_k)^2}. \tag{11}$$

Pour $n = 1$, 2 et 3, la formule (11) exprime la distance habituelle entre les points x et y.

Pour la métrique (10), $\|x\| = \|x - 0\|$ est la distance entre x et 0 ou bien le *module* (*longueur*) du vecteur x.

La formule (7) signifie donc que pour la norme engendrée par le produit scalaire on a le **théorème des diagonales du parallélogramme** (fig. 1) : *la somme des carrés des diagonales du parallélogramme est égale à la somme des carrés de ses côtés.*

Pour la métrique euclidienne (11) l'inégalité triangulaire (9) prend la forme

$$\sqrt{\sum_{k=1}^{n}(x_k - y_k)^2} \leqslant \sqrt{\sum_{k=1}^{n}(x_k - z_k)^2} + \sqrt{\sum_{k=1}^{n}(z_k - y_k)^2}.$$

En posant ici $x_k - z_k = u_k$ et $z_k - y_k = v_k$, nous obtenons l'inégalité

$$\sqrt{\sum_{k=1}^{n}(u_k + v_k)^2} \leqslant \sqrt{\sum_{k=1}^{n}u_k^2} + \sqrt{\sum_{k=1}^{n}v_k^2}$$

(qui est vérifiée pour tout $(u, v) \in \mathbf{R}^n \times \mathbf{R}^n$). Cette inégalité s'appelle *inégalité de Minkowski.*

DÉFINITION 4. L'espace $\mathbf{R}^n$ muni du produit scalaire standard, de la norme euclidienne et de la métrique euclidienne s'appelle *espace euclidien de dimension n.*

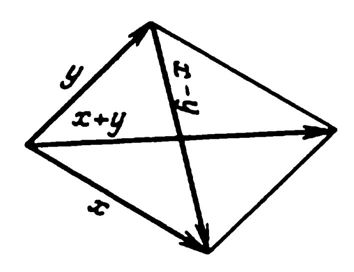


Fig. 1

6. Angles. Orthogonalité *)

Le produit scalaire permet d'introduire non seulement les distances mais les *angles* aussi. Soit X un espace vectoriel réel muni d'un produit scalaire et de la norme (6) que ce produit engendre, et soit (x, y) un couple quelconque de vecteurs non nuls de X (si bien que $\|x\| \neq 0$, $\|y\| \neq 0$). En vertu de l'inégalité de Cauchy-Bouniakovski, $|\langle x, y\rangle| \leqslant \|x\|\,\|y\|$, d'où

$$\left|\frac{\langle x, y\rangle}{\|x\|\,\|y\|}\right| \leqslant 1.$$

Or tout nombre dont le module ne dépasse pas 1 est le cosinus d'un angle bien défini $\varphi \in [0, \pi]$. Donc,

$$\frac{\langle x, y\rangle}{\|x\|\,\|y\|} = \cos \varphi.$$

On dit que φ est l'*angle des vecteurs x et y*. Il en résulte que

$$\langle x, y\rangle = \|x\|\,\|y\| \cos \varphi,$$

i.e. *le produit scalaire de deux vecteurs non nuls est égal au produit des longueurs de ces vecteurs par le cosinus de l'angle qu'ils forment.* Cette affirmation est également valable en algèbre vectorielle sur le plan et dans l'espace où elle joue le rôle de la définition du produit scalaire, tandis qu'ici elle représente une conséquence de l'inégalité de Cauchy-Bouniakovski.

DÉFINITION 5. Les vecteurs x, y sont dits *orthogonaux* si leur produit scalaire $\langle x, y\rangle$ est nul.

Si x et y sont des vecteurs non nuls, leur orthogonalité signifie que l'angle formé par ces vecteurs est droit (d'où le terme « orthogonalité »). Le vecteur nul est orthogonal à tout vecteur.

En particulier, *les fonctions réelles $x(t)$ et $y(t)$ continues sur un intervalle $[a, b]$ sont dites orthogonales* (sur cet invervalle) *si* $\int_a^b x(t)\, y(t)\, dt = 0$. La notion d'orthogonalité des fonctions joue un rôle important dans la théorie des séries de Fourier.

§ 2. Espaces métriques

1. Notion d'espace métrique

DÉFINITION 1. On appelle *espace métrique* un ensemble non vide muni d'une métrique.

---

*) Nous aurons besoin de ce numéro dans le chapitre 6.

Chaque ensemble non vide de points de la droite devient un espace métrique si on le munit, en qualité de métrique, d'une distance habituelle. Mais il existe sur la droite une infinité d'autres métriques. Ainsi chaque fonction injective f (par exemple, une fonction strictement monotone) engendre sur la droite la métrique

$$d(x, y) = |f(x) - f(y)|.$$

En effet, $|f(x) - f(y)| \leqslant |f(x) - f(z)| + |f(y) - f(z)|$ et, si $|f(x) - f(y)| = 0$, $x = y$ en vertu de l'injectivité de la fonction f. Pour $f(x) = \mathrm{Arctg}\, x$, la distance de deux points quelconques de la droite est inférieure à π. S'il s'agit du cercle, il est plus naturel de prendre pour la métrique la « distance curviligne » et non pas la distance habituelle, i.e. pour tous points A, B du cercle on prend pour $d(A, B)$ la moindre des longueurs de deux arcs de cercle reliant A et B. Par la suite nous discuterons des exemples plus compliqués d'espaces métriques.

DÉFINITION 2. Soient X un espace métrique et d la métrique de X. On dit que la suite (x_n) de points de X *converge vers le point* $x \in X$, ou bien que x est la *limite* de la suite (x_n), et on note $x_n \to x$, ou $x = \lim x_n$, si $d(x_n, x) \to 0$.

EXERCICE 1. Montrer que la convergence de la suite de points de la droite numérique pour la métrique $d(x, y) = |\mathrm{Arctg}\, x - \mathrm{Arctg}\, y|$ est équivalente à la convergence ordinaire, i.e. $d(x_n, x) \to 0 \Leftrightarrow x_n \to x$ dans un sens habituel.

La convergence des suites dans un espace métrique possède des propriétés habituelles :

1° *Si $x_n \to x$ et $x_n \to y$, alors $x = y$* (unicité de la limite). En effet, $0 \leqslant d(x, y) \leqslant d(x, x_n) + d(y, x_n)$ pour chaque $n \in \mathbf{N}$. Mais $d(x, x_n) + d(y, x_n)$ tend vers zéro. Par conséquent, $d(x, y) = 0$, i.e. $x = y$.

2° *Si x_n est égal à x pour tout n suffisamment grand*, $x_n \to x$. En effet, $d(x_n, x)$ est égal à zéro pour tout n suffisamment grand et, par conséquent, tend vers zéro.

3° *Si la suite (x_n) converge vers x, toute sous-suite (x_{n_k}) de (x_n) converge vers x.* En effet, $(d(x_{n_k}, x))$ est une sous-suite de la suite numérique $(d(x_n, x))$ et, par conséquent, tend aussi vers zéro.

EXERCICE 2. Montrer que $x_n \to x$ *si toute suite extraite de (x_n) possède une sous-suite qui converge vers x.*

Signalons encore deux inégalités utiles qui sont valables pour la métrique.

Tout triplet (x, y, z) de points d'un espace métrique muni de la métrique d vérifie l'inégalité

$$|d(x, y) - d(x, z)| \leqslant d(y, z) \tag{1}$$

(« la différence des longueurs de deux côtés du triangle ne dépasse pas la longueur du troisième côté »). En effet, $d(x, y) \leqslant d(x, z) + d(y, z)$ d'après l'inégalité triangulaire, d'où $d(x, y) - d(x, z) \leqslant d(y, z)$. En remplaçant y par z et *vice versa*, nous obtenons aussi $d(x, z) - d(x, y) \leqslant d(z, y) = d(y, z)$. Par conséquent, $|d(x, y) - d(x, z)| = \max(d(x, y) - d(x, z), d(x, z) - d(x, y)) \leqslant d(y, z)$.

Tout quadruplet (x, y, z, t) *de points d'un espace métrique muni de la métrique d satisfait à l'inégalité*

$$|d(x, y) - d(z, t)| \leqslant d(x, z) + d(y, t). \tag{2}$$

En effet, en vertu de (1)

$$|d(x, y) - d(z, t)| \leqslant |d(x, y) - d(y, z)| + |d(y, z) - d(z, t)| = \\ = |d(y, x) - d(y, z)| + |d(z, y) - d(z, t)| \leqslant d(x, z) + d(y, t).$$

2. Notion d'espace normé

DÉFINITION 3. On appelle *espace normé réel* un espace vectoriel réel muni d'une norme et de la métrique engendrée par cette norme.

EXEMPLE 1. Conformément à la définition 4 du § 1, l'espace euclidien n-dimensionnel $\mathbf{R}^n$ est un espace normé réel. Sa norme est

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2} \tag{3}$$

et sa métrique est

$$d(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} = \|x - y\|. \tag{4}$$

La convergence des suites dans $\mathbf{R}^n$ est définie en coordonnées, i.e. $x^{(m)} \to a$ si et seulement si $x_k^{(m)} \to a_k$ pour $k = 1, \ldots, n$, ce qui découle immédiatement des inégalités

$$0 \leqslant \max_{1 \leqslant k \leqslant n} |x_k^{(m)} - a_k| \leqslant d(x^{(m)}, a) \leqslant \\ \leqslant |x_1^{(m)} - a_1| + \ldots + |x_n^{(m)} - a_n|.$$

EXEMPLE 2. *Espace hilbertien* l_2. C'est une généralisation des espaces euclidiens $\mathbf{R}^n$ en dimension infinie. Les éléments de l'espace l_2 sont toutes les suites $x = (x_n)$ de nombres réels dont les séries $\sum x_n^2$ des carrés des « coordonnées » sont convergentes. L'introduction dans l_2 de l'addition, de la multiplication par un nombre et du produit scalaire est fondée sur les identités numériques suivantes :

$$(a + b)^2 + (a - b)^2 = 2(a^2 + b^2), \tag{5}$$

$$(6) \qquad ab = \tfrac{1}{4}[(a + b)^2 - (a - b)^2].$$

Rappelons que si $x = (x_n)$ et $y = (y_n)$ sont des suites numériques, on entend par $x + y$ et $x - y$ les suites respectives $(x_n + y_n)$ et $(x_n - y_n)$ et par λx, la suite (λx_n).

Si $x, y \in l_2$, il en est de même pour $x \pm y$. En effet, les séries positives $\sum(x_n + y_n)^2$ et $\sum(x_n - y_n)^2$ sont majorées en vertu de l'identité (5) par une série positive convergente $\sum 2(x_n^2 + y_n^2)$ et, par conséquent, convergent.

Si $x \in l_2$, alors $\lambda x \in l_2$ pour tout $\lambda \in \mathbf{R}$. En effet, la série $\sum(\lambda x_n)^2 = \lambda^2 \sum x_n^2$ converge avec $\sum x_n^2$.

Ainsi, dans l_2 sont définies les opérations d'addition et de multiplication de ses éléments par des nombres réels. Il est clair que l_2 est un espace vectoriel réel pour ces opérations.

Il résulte de la convergence des séries $\sum(x_n + y_n)^2$ et $\sum(x_n - y_n)^2$ en vertu de l'identité (6) que *la série $\sum x_n y_n$ converge pour tous les $x, y \in l_2$.* Posons

$$(7) \qquad \langle x, y \rangle = \sum_{n=1}^{\infty} x_n y_n.$$

On voit aisément que la fonction $\langle x, y \rangle$ ainsi définie sur $l_2 \times l_2$ satisfait aux axiomes du produit scalaire. On la prend pour le *produit scalaire standard sur l_2*. L'espace vectoriel l_2 muni de ce produit scalaire ainsi que de la norme

$$\|x\| = \sqrt{\sum_{n=1}^{\infty} x_n^2}$$

et de la métrique

$$d(x, y) = \sqrt{\sum_{n=1}^{\infty} (x_n - y_n)^2} = \|x - y\|$$

que ce produit engendre a été étudié pour la première fois par Hilbert et porte en son honneur le nom d'*espace hilbertien*.

Comme $|x_k^{(n)} - x_k| \leqslant d(x^{(n)}, x)$ pour tout $k \in \mathbf{N}$, la convergence de $x^{(n)}$ vers x entraîne celle de $x_k^{(n)}$ vers x_k pour tout k, i.e. la convergence pour la métrique de l'espace l_2 entraîne la convergence en coordonnées. Mais la réciproque n'est pas vraie. En effet, soit $e^{(n)}$ le vecteur de l_2 dont la n-ième coordonnée est égale à 1 et toutes les autres sont nulles. Il est évident que $e^{(n)}$ converge en coordonnées vers le vecteur nul 0 (i.e. vers le vecteur dont toutes les coordonnées sont égales à zéro). Cependant, $\|e^{(n)} - 0\| = \|e^{(n)}\| = 1$, de sorte que $e^{(n)}$ ne converge pas pour la métrique de l_2 vers le vecteur 0.

EXEMPLE 3. *Espace* $C(\Delta)$. Ses éléments sont toutes les fonctions réelles continues sur un intervalle fermé borné Δ. Elles forment un espace vectoriel réel pour les opérations ordinaires d'addition des fonctions et de multiplication d'une fonction par un nombre réel. En outre, le module de toute fonction continue est continu et, par conséquent, présente un maximum sur l'intervalle Δ. Montrons que $\max_{t\in\Delta} |x(t)|$ est une *norme sur* $C(\Delta)$.

En effet :

1) Vu que pour tout $t \in \Delta$

$$|x(t) + y(t)| \leqslant |x(t)| + |y(t)| \leqslant \max_{t\in\Delta} |x(t)| + \max_{t\in\Delta} |y(t)|,$$

on a $\max_{t\in\Delta} |x(t) + y(t)| \leqslant \max_{t\in\Delta} |x(t)| + \max_{t\in\Delta} |y(t)|$;

2) $\max_{t\in\Delta} |\lambda x(t)| = |\lambda| \max_{t\in\Delta} |x(t)|$;

3) si $x \neq 0$, i.e. $x(t) \not\equiv 0$, $\max_{t\in\Delta} |x(t)| \neq 0$.

EXERCICE 3. Montrer que cette norme n'est pas engendrée par un produit scalaire (voir exercice 2 du § 1).

$C(\Delta)$ est donc un espace réel normé. *La convergence des suites dans* $C(\Delta)$ *est une convergence uniforme sur* Δ. En effet, $x_n \to x$ dans $C(\Delta)$ signifie que $\max_{t\in\Delta} |x_n(t) - x(t)| \to 0$, i.e.

$$(\forall \varepsilon > 0)(\exists N)(\forall n \in \mathbf{N})(n \geqslant N \Rightarrow \max_{t\in\Delta} |x_n(t) - x(t)| \leqslant \varepsilon),$$

ce qui est équivalent à la condition

$$(\forall \varepsilon > 0)(\exists N)(\forall n \in \mathbf{N})(\forall t \in \Delta)(n \geqslant N \Rightarrow |x_n(t) - x(t)| \leqslant \varepsilon),$$

i.e. à la convergence uniforme de $x_n(t)$ vers $x(t)$ sur Δ.

§ 3. Voisinages d'un point dans un espace métrique

1. Boule ouverte ; ε-voisinage

DÉFINITION 1. Soit X un espace métrique muni d'une métrique d et soit $r > 0$. On appelle *boule ouverte dans* X *de centre* a *et de rayon* r l'ensemble

(1) $$U(a\,;r) = \{x \in X \mid d(x, a) < r\}.$$

Notons que $a \in U(a\,;r)$ (car $d(a, a) = 0 < r$) et que $U(a\,;r_1) \subset U(a\,;r_2)$ si $0 < r_1 < r_2$ (car $d(x, a) < r_1$ entraîne $d(x, a) < r_2$).

Dans $\mathbf{R}^n$ muni de la métrique standard,

(2) $$U(a\,;r) = \left\{x \in \mathbf{R}^n \,\middle|\, \sum_{k=1}^{n}(x_k - a_k)^2 < r^2\right\}.$$

En particulier, $U(a\,;r)$ est un intervalle $]a - r, a + r[$ pour $n = 1$, $U((a, b)\,;r)$ est un disque ouvert

$$\{(x, y) \in \mathbf{R}^2 \mid (x - a)^2 + (y - b)^2 < r^2\}$$

pour $n = 2$, $U((a, b, c)\,;r)$ est une boule ouverte

$$\{(x, y, z) \in \mathbf{R}^3 \mid (x - a)^2 + (y - b)^2 + (z - c)^2 < r^2\}$$

pour $n = 3$ (d'où la dénomination générale de « boule »).

DÉFINITION 2. On appelle *ε-voisinage* d'un point a de l'espace métrique X la boule ouverte $U(a\,;\varepsilon)$.

En particulier, dans $\mathbf{R}$ muni d'une distance habituelle, le ε-voisinage $U(a\,;\varepsilon)$ du nombre a est encore un intervalle $]a - \varepsilon, a + \varepsilon[$ de rayon ε et de centre a. Dans $\mathbf{R}^2$, le ε-voisinage $U((a, b)\,;\varepsilon)$ du point (a, b) est un disque ouvert de rayon ε et de centre (a, b).

2. Notion générale de voisinage

Même pour les études dans le plan, il est rationnel d'élargir la notion de *voisinage d'un point.*

DÉFINITION 3. On appelle *voisinage d'un point* de l'espace métrique X tout ensemble dans X contenant un ε-voisinage quelconque de ce point.

Ainsi, *U est un voisinage du point a dans X si et seulement si $U \subset X$ et s'il existe un $\varepsilon > 0$ tel que $U \supset U(a\,;\varepsilon)$*, i.e. *$U$ contient une boule ouverte de centre a.* En particulier, *tout ε-voisinage du point a est son voisinage au sens de la définition* 3. Le voisinage du point $a \in X$ peut aussi être représenté par une *boule fermée*

$$B(a\,;r) = \{x \in X \mid d(x, a) \leqslant r\}$$

de centre a et de rayon $r > 0$ car celle-ci contient $U(a\,;r)$.

Dans $\mathbf{R}^n$ on recourt souvent aussi à des *voisinages cubiques.*

DÉFINITION 4. On appelle *cube ouvert dans $\mathbf{R}^n$ de centre a et d'arête* $2h$ (où $h > 0$) l'ensemble

$$Q(a\,;h) = \{x \in \mathbf{R}^n \mid |x_k - a_k| < h \ (k = 1, \dots, n)\}. \tag{3}$$

Pour $n = 1$ c'est un intervalle $]a - h, a + h[$. Pour $n = 2$, l'ensemble $Q((a, b)\,;h)$ est un *carré*

$$\{(x, y) \in \mathbf{R}^2 \mid a - h < x < a + h, \ b - h < y < b + h\}$$

(fig. 2). Pour $n = 3$, l'ensemble $Q((a, b, c)\,;h)$ est un *cube*

$$\{(x, y, z) \in \mathbf{R}^3 \mid a - h < x < a + h, \ b - h < y < b + h, \ c - h < z < c + h\}$$

(fig. 3). (D'où la dénomination générale de « cube ».)

Chaque cube $Q(a\,;\varepsilon)$ dans $\mathbf{R}^n$ est un voisinage du point a car

$$Q(a\,;\varepsilon) \supset U(a\,;\varepsilon)$$

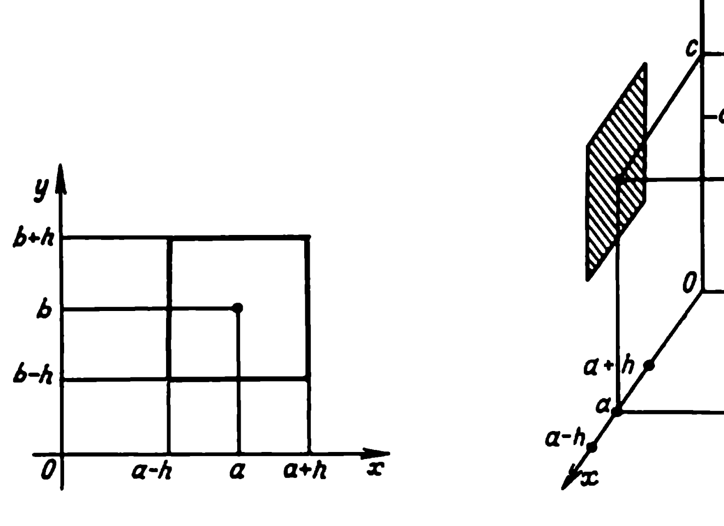


Fig. 2

Fig. 3

(fig. 4). En effet, vu que

$$|x_k - a_k| \leqslant \|x - a\| = \sqrt{\sum_{k=1}^{n}(x_k - a_k)^2} \quad (k = 1, \ldots, n),$$

il découle de $\|x - a\| < \varepsilon$ que $|x_k - a_k| < \varepsilon$ pour $k = 1, \ldots, n$, i.e. $x \in U(a\,;\varepsilon)$ entraîne $x \in Q(a\,;\varepsilon)$. Nous dirons que les ensembles $Q(a\,;\varepsilon)$ sont des *voisinages cubiques* du point a (et que les ensembles $U(a\,;\varepsilon)$ sont ses *voisinages sphériques*).

THÉORÈME 1. *La boule ouverte $U(a\,;r)$ est un voisinage de chacun de ses points.*

Démonstration. Soit $x \in U(a\,;r)$, i.e. $d(x, a) < r$, de façon que $\varepsilon = r - d(x, a) > 0$ (fig. 5). Si $y \in U(x\,;\varepsilon)$, i.e. $d(y, x) < \varepsilon$, il en résulte que

$$d(y, a) \leqslant d(y, x) + d(x, a) < \varepsilon + d(x, a) = r,$$

i.e. $y \in U(a\,;r)$. Donc, $U(x\,;\varepsilon) \subset U(a\,;r)$, de sorte que $U(a\,;r)$ est un voisinage de son point (quelconque) x.

EXERCICE 1. Montrer que $Q(a\,;h)$ est aussi un voisinage de l'un quelconque de ses points.

3. Propriétés fondamentales des voisinages

Notons $\mathscr{V}(a)$ l'ensemble de tous les voisinages du point a dans un espace métrique.

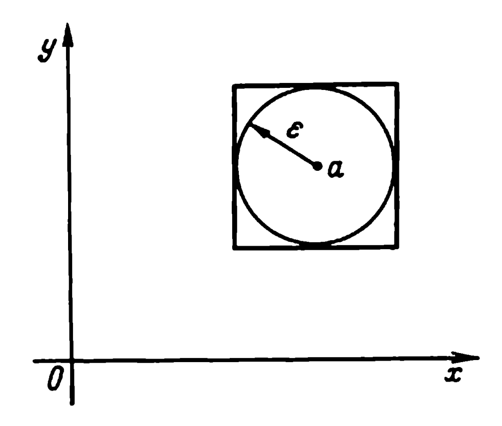


Fig. 4

THÉORÈME 2. *Les voisinages des points de l'espace métrique X possèdent les propriétés suivantes* :

1) *si U est un voisinage du point* $a \in X$ *et* $U \subset V \subset X$, *alors V est aussi un voisinage du point a* ;

2) *l'intersection de deux voisinages quelconques du point a est un voisinage de ce point* ;

3) *le point a appartient à chacun de ses voisinages* ;

4) *tout voisinage du point a contient un voisinage de ce point qui est voisinage de chacun de ses points.*

Démonstration. 1) Soient $U \in \mathscr{U}(a)$ et $U \subset V \subset X$. En vertu de la définition 3, il existe un $\varepsilon > 0$ tel que $U(a\,;\varepsilon) \subset U$, ce qui signifie à plus forte raison que $U(a\,;\varepsilon) \subset V$ et, par conséquent, $V \in \mathscr{U}(a)$.

2) Soient $U \in \mathscr{U}(a)$ et $V \in \mathscr{U}(a)$. Il existe donc des $\varepsilon_1 > 0$ et $\varepsilon_2 > 0$ tels que $U \supset U(a\,;\varepsilon_1)$ et $V \supset U(a\,;\varepsilon_2)$. Posons $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$. Alors $\varepsilon > 0$ et $U \cap V \supset U(a\,;\varepsilon_1) \cap U(a\,;\varepsilon_2) \supset U(a\,;\varepsilon)$, de sorte que $U \cap V \in \mathscr{U}(a)$.

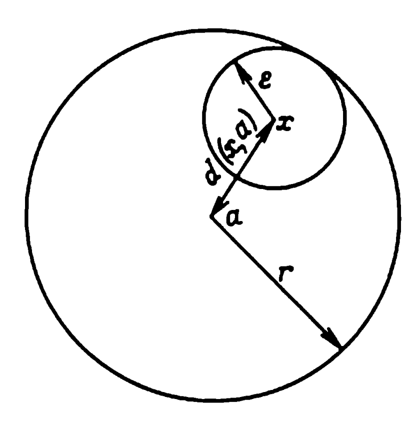


Fig. 5

3) Soit $U \in \mathscr{U}(a)$. Il existe donc un $\varepsilon > 0$ tel que $U(a\,;\varepsilon) \subset U$. Comme $a \in U(a\,;\varepsilon)$, on a à plus forte raison $a \in U$.

4) Soit $U \in \mathscr{U}(a)$. Il existe donc un $\varepsilon > 0$ tel que $U(a\,;\varepsilon) \subset U$. Mais $U(a\,;\varepsilon)$ est un voisinage du point a. Donc, en vertu du théorème 1, il est un voisinage de chacun de ses points.

§ 4. Espaces topologiques

1. Notion d'espace topologique

DÉFINITION 1. On dit que l'ensemble X est muni d'une *topologie* si à tout point de X est associée une famille non vide de sous-ensembles de X, appelés *voisinages* de ce point, satisfaisant aux conditions suivantes (*axiomes des voisinages*) :

1) tout sous-ensemble de X qui contient un voisinage du point est aussi un voisinage de ce point ;

2) l'intersection de deux voisinages quelconques du point est voisinage de ce point ;

3) le point appartient à chacun de ses voisinages ;

4) tout voisinage du point contient un voisinage de ce point qui est voisinage de chacun de ses points.

L'ensemble X muni d'une topologie donnée est appelé *espace topologique*. Deux topologies sur X se confondent si les familles de voisinages associées à ces topologies coïncident. L'ensemble vide est un espace topologique par définition ; il représente un voisinage de chacun de ses points en l'absence de ces derniers.

DÉFINITION 2. On appelle *ensemble ouvert* d'un espace topologique X tout sous-ensemble de X qui est voisinage de chacun de ses points.

Pour abréger on peut formuler l'axiome 4 des voisinages comme suit : tout voisinage du point contient un voisinage ouvert de ce point.

En vertu du théorème 2 du § 3, les voisinages des points de l'espace métrique satisfont aux axiomes 1 à 4 et lui confèrent par là même une topologie. Nous admettrons que l'espace métrique est muni de cette topologie et, par conséquent, nous traiterons chaque espace métrique en tant qu'espace topologique. Le théorème 1 du § 3 signifie que conformément à sa dénomination *chaque boule ouverte dans un espace métrique est un ensemble ouvert*.

Les espaces métriques ne sont pas les seuls espaces topologiques dont nous aurons besoin.

Introduisons une topologie dans $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$. Le voisinage de tout point $a \in \overline{\mathbf{R}}$ distinct de $\pm\infty$ est par définition chaque ensemble $U \subset \overline{\mathbf{R}}$ qui contient un ε-voisinage $]a - \varepsilon, a + \varepsilon[$ de ce point. Le voisinage de $+\infty$ (resp. de $-\infty$) dans $\overline{\mathbf{R}}$ est tout ensemble $U \subset \overline{\mathbf{R}}$ qui contient un intervalle

$]M, +\infty]$ (resp. $[-\infty, -M[$) où $M > 0$. Les voisinages des points $a \neq \pm\infty$ dans $\bar{\mathbf{R}}$ se distinguent de ceux dans $\mathbf{R}$ par la possibilité d'ajouter un ou deux points $+\infty$, $-\infty$, ce qui ne viole évidemment pas les axiomes des voisinages. Vérifions les axiomes des voisinages pour les points $\pm\infty$.

1. L'axiome 1 est vérifié par définition.

2. Si U_1 et U_2 sont des voisinages du point $+\infty$, il existe des nombres $M_1 > 0$ et $M_2 > 0$ tels que $U_1 \supset]M_1, +\infty]$ et $U_2 \supset]M_2, +\infty]$; mais alors $U_1 \cap U_2 \supset]M, +\infty]$ où $M = \max(M_1, M_2)$, de sorte que $U_1 \cap U_2$ est aussi un voisinage du point $+\infty$. Pour $-\infty$ la démonstration de l'axiome 2 est analogue.

3. Les points $+\infty$ et $-\infty$ sont inclus par définition dans chacun de leurs voisinages.

4. Par définition, tout voisinage U du point $+\infty$ contient un voisinage $]M, +\infty]$ où $M > 0$; mais ce dernier est un voisinage de chacun de ses points : il l'est du point $+\infty$ par hypothèse, et si $a \in]M, +\infty]$, alors $]M, +\infty] \supset]a - \varepsilon, a + \varepsilon]$, où $\varepsilon = a - M$. Pour $-\infty$ la démonstration de l'axiome 4 est analogue.

EXERCICE 1. Montrer que la métrique ordinaire $d(x, y) = |x - y|$ sur $\mathbf{R}$ ne peut pas être prolongée jusqu'à une métrique sur $\bar{\mathbf{R}}$ de telle sorte que cette dernière engendre la topologie décrite sur l'espace $\bar{\mathbf{R}}$, i.e. qu'un ensemble dans $\bar{\mathbf{R}}$ soit le voisinage du point $a \in \bar{\mathbf{R}}$ si et seulement s'il contient un ε-voisinage de ce point (dans la métrique prolongée).

EXERCICE 2. Posons $\operatorname{Arctg}(+\infty) = \pi/2$, $\operatorname{Arctg}(-\infty) = -\pi/2$ et $d(x, y) = |\operatorname{Arctg} x - \operatorname{Arctg} y|$ pour tous $x, y \in \bar{\mathbf{R}}$. En déduire que d est une métrique sur $\bar{\mathbf{R}}$ et que la topologie de l'espace $\bar{\mathbf{R}}$ est engendrée par cette métrique (de sorte que l'espace $\bar{\mathbf{R}}$ est dit *métrisable*).

Introduisons une topologie dans l'ensemble $\widehat{\mathbf{R}^n}$ qu'on obtient par adjonction à $\mathbf{R}^n$ d'un seul point ∞ « à l'infini ». Appelons voisinage d'un point $a \in \mathbf{R}^n$ dans $\widehat{\mathbf{R}^n}$ tout ensemble $U \subset \widehat{\mathbf{R}^n}$ qui contient une boule ouverte $U(a ; \varepsilon)$, et voisinage du point ∞, tout ensemble $U \subset \widehat{\mathbf{R}^n}$ qui contient dans $\widehat{\mathbf{R}^n}$ le complémentaire d'une boule fermée $B(O ; M)$, i.e. tout ensemble qui contient le point ∞ et tous les points $x \in \mathbf{R}^n$ pour lesquels $\|x\| > M$.

EXERCICE 3. Vérifier les axiomes des voisinages dans $\widehat{\mathbf{R}^n}$.

EXERCICE 4. Montrer que tout comme dans le cas de $\bar{\mathbf{R}}$, on ne peut pas prolonger la métrique euclidienne sur $\mathbf{R}^n$ jusqu'à une métrique sur $\widehat{\mathbf{R}^n}$ de telle façon que cette dernière engendre une topologie de l'espace $\widehat{\mathbf{R}^n}$.

EXERCICE 5. Montrer que $\widehat{\mathbf{R}^n}$ est quand même métrisable.

Citons encore un exemple d'espace topologique.

EXEMPLE 1. Soit X une famille non vide de fonctions réelles définies sur un ensemble non vide E. Pour toute fonction $f_0 \in X$, tout ensemble fini de points $x_1, \ldots, x_n \in E$ et tout $\varepsilon > 0$ posons

$$U_{x_1, \ldots, x_n; \varepsilon}(f_0) = \{f \in X \mid |f(x_k) - f_0(x_k)| < \varepsilon \ (k = 1, \ldots, n)\}.$$

Appelons *voisinage de la fonction* f_0 dans X tout ensemble de fonctions de X qui contient l'ensemble de la forme $U_{x_1, \ldots, x_n; \varepsilon}(f_0)$. On voit aisément que les axiomes des voisinages sont vérifiés.

REMARQUE. Pour $X = C(\Delta)$ (exemple 3 du § 2) et $E = \Delta$, la topologie décrite n'est pas métrisable.

2. Système fondamental de voisinages

Les voisinages sphériques et cubiques dans $\mathbf{R}^n$ sont équivalents dans un certain sens.

DÉFINITION 3. On appelle *système fondamental de voisinages* du point d'un espace topologique (et métrique, en particulier) toute famille de voisinages de ce point possédant la propriété suivante : tout voisinage du point contient un voisinage qui appartient à cette famille.

Ainsi, *l'ensemble de tous les voisinages sphériques $U(a\,;\varepsilon)$ du point a de l'espace métrique est un système fondamental de voisinages de ce point* car, par définition, tout voisinage du point a contient un voisinage sphérique de a. *Les voisinages $U(a\,;1/k)$ ($k \in$ N) forment* eux aussi *un système fondamental de voisinages du point a.* En effet, pour tout voisinage U du point a il existe un $\varepsilon > 0$ tel que $U \supset U(a\,;\varepsilon)$, et pour ce ε, un $k \in \mathbf{N}$ tel que $\varepsilon > 1/k$; mais alors $U \supset U(a\,;1/k)$. Ainsi donc, *dans un espace métrique chaque point possède une* suite *fondamentale de voisinages.*

Les voisinages cubiques du point dans $\mathbf{R}^n$ forment aussi un système fondamental de voisinages de ce point. En effet, nous avons déjà vu que les ensembles $Q(a\,;\varepsilon)$ sont des voisinages du point a. Or *tout voisinage sphérique du point a dans $\mathbf{R}^n$ contient un voisinage cubique*, à savoir

$$U(a\,;\varepsilon) \supset Q\left(a\,;\frac{\varepsilon}{\sqrt{n}}\right)$$

(fig. 6) : si $x \in Q\left(a\,;\frac{\varepsilon}{\sqrt{n}}\right)$, i.e. $|x_k - a_k| < \frac{\varepsilon}{\sqrt{n}}$, alors

$$\sum_{k=1}^{n}(x_k - a_k)^2 < \underbrace{\frac{\varepsilon^2}{n} + \ldots + \frac{\varepsilon^2}{n}}_{n \text{ fois}} = \varepsilon^2,$$

d'où $\|x - a\| < \varepsilon$, i.e. $x \in U(a\,;\varepsilon)$. Ainsi, pour tout voisinage U du point a il existe un $\varepsilon > 0$ tel que $U \supset U(a\,;\varepsilon)$ et, par conséquent, $U \supset Q(a\,;\varepsilon/\sqrt{n})$.

3. Convergence des suites

DÉFINITION 4. On dit que la suite (x_n) de points de l'espace topologique X *converge* vers le point $a \in X$, ou que a est la *limite* de la suite (x_n), et on écrit $x_n \to a$, ou bien $a = \lim x_n$, si x_n appartient à chaque voisinage du point a pour tout n suffisamment grand *).

---

*) Notons que si l'on prend pour les voisinages des points de l'espace X tous les ensembles de X qui contiennent ce point (topologie dite *discrète* sur X), la définition introduite de la limite d'une suite dit que seules les suites stationnaires ont une limite.

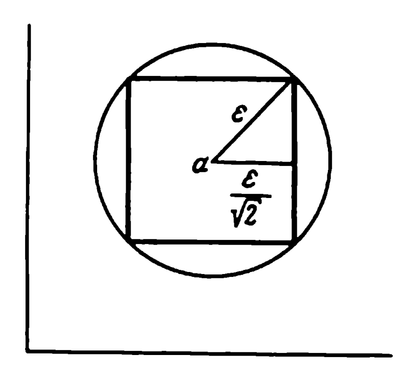


Fig. 6

Pour les espaces métriques, cette définition est équivalente à la définition 2 du § 2. En effet, conformément à cette dernière, $x_n \to a$ si $d(x_n, a) \to 0$, i.e. quel que soit $\varepsilon > 0$ on a $d(x_n, a) < \varepsilon$ pour tout n suffisamment grand, ou, ce qui revient au même, quel que soit $\varepsilon > 0$ on a $x_n \in U(a\,;\varepsilon)$ pour tout n suffisamment grand. De par la définition d'un voisinage dans un espace métrique (définition 3 du § 3), ceci équivaut à ce que x_n appartient à chaque voisinage du point a pour tout n suffisamment grand, i.e. $x_n \to a$ au sens de la définition 4.

Signalons maintenant qu'en général *une suite* (x_n) *de points de l'espace topologique* X *converge vers un point* $a \in X$ *si tout voisinage d'un système fondamental* $\mathcal{V}$ *de voisinages du point a contient* x_n *pour tout n suffisamment grand.* En effet, de par la définition du système $\mathcal{V}$, pour chaque voisinage U du point a il existe un $V \in \mathcal{V}$ contenu dans U ; si $x_n \in V$ pour tout n suffisamment grand, on a à plus forte raison $x_n \in U$, i.e. la condition de la définition 4 est vérifiée.

Comme les voisinages cubiques de tout point a forment dans $\mathbf{R}^n$ un système fondamental de voisinages de a, il s'ensuit en particulier que $x^{(m)} \to a$ dans $\mathbf{R}^n$ si et seulement si $x^{(m)}$ appartient à chaque voisinage cubique de a pour tout m suffisamment grand, i.e. pour tout $\varepsilon > 0$ il existe un N tel que

$$|x_k^{(m)} - a_k| < \varepsilon \text{ pour tous } m \geqslant N \text{ et } k = 1, \ldots, n,$$

ou, ce qui revient au même, $x_k^{(m)} \to a_k$ pour $k = 1, \ldots, n$. Donc, nous avons obtenu de nouveau que la convergence dans $\mathbf{R}^n$ est une convergence en coordonnées.

Chaque nombre $a \in \bar{\mathbf{R}}$ possède un système fondamental de voisinages que forment les intervalles $]a - \varepsilon, a + \varepsilon[$, de sorte que la convergence de la suite (x_n) vers a dans $\widehat{\mathbf{R}}$ est la même que dans $\mathbf{R}$. Le point $+\infty$ (resp. $-\infty$) possède un système fondamental de voisinages que forment les intervalles $]M, +\infty]$ (resp. $[-\infty, -M[$) où M sont des nombres strictement

positifs, de sorte que pour les suites numériques la convergence vers $+\infty$ ou $-\infty$ dans $\bar{\mathbf{R}}$ est la même que dans $\mathbf{R}$. D'une façon analogue, la convergence dans $\widehat{\mathbf{R}^n}$ de la suite $(x^{(m)})$ vers $a \neq \infty$ est la même que dans $\mathbf{R}^n$. D'autre part, *si* $(x^{(m)}) \subset \mathbf{R}^n$, *alors* $x^{(m)} \to \infty$ *si et seulement si* $\|x^{(m)}\| \to +\infty$. En effet,

$$\mathbf{R}^n \setminus B(0\,;M) = \{x \in \mathbf{R}^n \mid \|x\| > M\},$$

et $x^{(m)} \to \infty$ signifie que $x^{(m)} \in \mathbf{R}^n \setminus B(0\,;M)$ pour tout m suffisamment grand.

Dans l'espace X de l'exemple 1, $f_m \to f_0$ si et seulement si $f_m(x) \to f_0(x)$ pour chaque point $x \in E$. En effet, $f_m \to f_0$ implique que pour tout point $x \in E$ et tout $\varepsilon > 0$ la fonction f_m appartient à $U_{x\,;\,\varepsilon}$ pour tout m suffisamment grand, i.e. $|f_m(x) - f_0(x)| < \varepsilon$ pour tout m suffisamment grand, si bien que $f_m(x) \to f_0(x)$. Réciproquement : si $f_m(x) \to f_0(x)$ pour tout $x \in E$, il existe pour chaque famille finie de points $x_1, \ldots, x_n \in E$ et tout $\varepsilon > 0$ un N tel que

$$|f_m(x_k) - f_0(x_k)| < \varepsilon$$

quels que soient $m \geqslant N$ et $k = 1, \ldots, n$, i.e. f_m appartient à chaque voisinage $U_{x_1,\ldots,x_n;\varepsilon}(f_0)$ de f_0 pour tout m suffisamment grand ; or ces voisinages forment un système fondamental de voisinages de f_0. La topologie de l'espace X est appelée *topologie de la convergence simple.*

4. Adhérence et intérieur d'un ensemble

Les voisinages permettent de définir deux relations topologiques de proximité d'un point et d'un ensemble.

DÉFINITION 5. Soient X un espace topologique, $a \in X$ et $E \subset X$. Le point a est appelé *point adhérent* à l'ensemble E si tout voisinage du point a a une intersection non vide avec E. L'ensemble des points adhérents à E, noté $\bar{E}$ ou $[E]$, s'appelle *adhérence* de l'ensemble E. Le point a s'appelle *point intérieur* à l'ensemble E si E est un voisinage du point a. L'ensemble des points intérieurs à E, noté $\dot{E}$ ou $]E[$, s'appelle *intérieur* de l'ensemble E.

1° *Si X est un espace topologique, $E \subset X$ et a est la limite d'une suite de points de E, alors $a \in \bar{E}$.* En effet, si $a = \lim x_n$ où $x_n \in E$, la définition 4 dit que tout voisinage du point a contient des éléments de la suite (x_n) et, par conséquent, a une intersection non vide avec E.

Les notions d'adhérence et d'intérieur d'un ensemble sont intimement liées. Désignons le complémentaire $X \setminus E$ de l'ensemble E dans X par $\complement_X E$ et rappelons ses propriétés :

$$E \cup \complement_X E = X,\ E \cap \complement_X E = \varnothing,\ \complement_X(\complement_X E) = E,\ E \subset F \Leftrightarrow \complement_X E \supset \complement_X F.$$

2° *Tout ensemble E d'un espace topologique X vérifie les relations suivantes :*

a) $\complement_X[E] =]\complement_X E[$, b) $]E[= \complement_X[\complement_X E]$,
c) $[E] = \complement_X]\complement_X E[$, d) $\complement_X]E[= [\complement_X E]$.

En effet, $x \in \complement_X[E]$ signifie que x n'est pas un point adhérent à E, i.e. il existe un voisinage U du point x qui ne se coupe pas avec E ou, ce qui revient au même, qui est contenu dans $\complement_X E$. Mais en vertu de l'axiome 1 des voisinages, cela signifie que $\complement_X E$ est un voisinage du point x, i.e. $x \in]\complement_X E[$. Donc, les relations $x \in \complement_X[E]$ et $x \in]\complement_X E[$ sont équivalentes, i.e. la relation a) est juste. La relation b) résulte de a) par substitution de $\complement_X E$ à E, et les relations c) et d) par application respectivement à a) et b) de l'opération de passage au complémentaire.

Les propriétés fondamentales de l'opération $E \to [E]$ de passage à l'adhérence d'un ensemble sont formulées dans quatre propositions suivantes.

3° $E \subset [E]$. En effet, soit $a \in E$. En vertu de l'axiome 3 des voisinages, tout voisinage du point a contient a et, par suite, un point de E. Donc, $a \in [E]$.

4° *Si* $E \subset F \subset X$, *alors* $[E] \subset [F]$. En effet, les voisinages du point adhérent à l'ensemble E se coupent avec E et, par suite, se coupent avec tout ensemble $F \subset X$ qui contient E.

5° $[[E]] = [E]$. Soient $x \in [[E]]$ et U un voisinage arbitraire du point x. En vertu de l'axiome 4 des voisinages, U contient un voisinage ouvert V du point x. Comme x est un point adhérent à $[E]$ et V est son voisinage, il vient $V \cap [E] \neq \varnothing$. Soit $y \in V \cap [E]$. Comme $y \in [E]$ et V est un voisinage de son point y parce qu'un ensemble ouvert, V et à plus forte raison U se coupent avec E. Ainsi, tout voisinage du point x a une intersection non vide avec E, i.e. $x \in [E]$. Donc, $x \in [[E]]$ entraîne $x \in [E]$, i.e. $[[E]] \subset [E]$. L'inclusion inverse découle de 3°.

6° $[E \cup F] = [E] \cup [F]$. En effet, comme $E \subset E \cup F$ et $F \subset E \cup F$, on a $[E] \cup [F] \subset [E \cup F]$ en vertu de 4°. Soit maintenant $x \notin [E] \cup [F]$. Vu que $x \notin [E]$, il existe un voisinage U_1 du point x qui ne se coupe pas avec E. De façon analogue, puisque $x \notin [F]$, le point x possède un voisinage U_2 qui ne se coupe pas avec F. Mais alors $U_1 \cap U_2$ est un voisinage du point x (axiome 2 des voisinages) qui ne se coupe pas avec $E \cup F$, de sorte que $x \notin [E \cup F]$. Donc, $x \notin [E] \cup [F]$ entraîne $x \notin [E \cup F]$, i.e. $\complement_X([E] \cup \cup [F]) \subset \complement_X[E \cup F]$, d'où découle $[E \cup F] \subset [E] \cup [F]$.

Chaque proposition 3° à 6° pour l'adhérence de l'ensemble entraîne, en vertu des relations 2°, une proposition « duale » pour l'intérieur de l'ensemble :

3*. $E \supset]E[$. En effet, en vertu de 3°, $\complement_X E \subset [\complement_X E]$. En passant aux complémentaires et en utilisant 2° a), nous obtenons

$$E = \complement_X(\complement_X E) \supset \complement_X[\complement_X E] =]\complement_X(\complement_X E)[=]E[.$$

4*. $E \subset F(\subset X) \Rightarrow \,]E[\, \subset \,]F[$. En vertu des propositions 4° et 2° d),

$$E \subset F \Leftrightarrow C_X F \subset C_X E \Rightarrow [C_X F] \subset [C_X E] \Leftrightarrow C_X]F[\subset C_X]E[\Leftrightarrow$$
$$\Leftrightarrow \,]E[\, \subset \,]F[.$$

5*. $]]E[[\, = \,]E[$. En vertu des propositions 2° b), 5° et 2° a),

$$]E[\, = C_X[C_X E] = C_X[[C_X E]] = \,]C_X[C_X E][\, = \,]]C_X(C_X E)[[\, = \,]]E[[.$$

EXERCICE 6. Formuler et démontrer la proposition 6* duale à 6°.

5. Ensemble fermé et ensemble ouvert

DÉFINITION 6. On appelle *ensemble fermé* d'un espace topologique X tout ensemble E de X qui contient tous ses points adhérents.

7° *Un ensemble fermé contient les limites de toutes les suites convergentes de ses points.* Cela découle directement de 1°.

EXEMPLE 2. *Les intervalles* $[a, b]$ *sont fermés dans* $\mathbf{R}$. Si $x \notin [a, b]$, alors $x \notin \overline{[a, b]}$ car x est séparé de $[a, b]$ par son $(a - x)$-voisinage s'il est inférieur à a, et par son $(x - b)$-voisinage s'il est supérieur à b, et, par conséquent, x n'est pas un point adhérent à $[a, b]$. Pour cette raison, $x \in \overline{[a, b]} \Rightarrow x \in [a, b]$, d'où $[a, b] = \overline{[a, b]}$. De façon analogue, on montre que les intervalles $[a, +\infty[$ et $]-\infty, b]$ sont fermés dans $\mathbf{R}$. Mais dans $\bar{\mathbf{R}}$ ils ne le sont plus parce qu'ils ne contiennent pas leurs points adhérents $+\infty$ et $-\infty$ respectivement. Les intervalles $[a, b]$, où $-\infty \leqslant a \leqslant b \leqslant +\infty$, sont fermés dans $\mathbf{R}$.

EXEMPLE 3. *La boule fermée* $B(a\,;r)$ (voir n° 2 du § 3) *dans un espace métrique* X *est un ensemble fermé* (d'où sa dénomination). Soit $x \in [B(a\,;r)]$. Il existe alors pour tout $\varepsilon > 0$ un point $x_\varepsilon \in U(a\,;\varepsilon) \cap B(a\,;r)$, d'où $d(x, a) \leqslant d(x, x_\varepsilon) + d(x_\varepsilon, a) < \varepsilon + r$. Comme ε est arbitraire, nous en tirons que $d(x, a) \leqslant r$, i.e. $x \in B(a\,;r)$.

8° *L'ensemble* E *d'un espace topologique est fermé si et seulement si* $E = [E]$. Conformément aux définitions 5 et 6, E est fermé si et seulement si $E \supset [E]$. Et en vertu de 3°, on a toujours $E \subset [E]$.

9° *L'adhérence* $[E]$ *de tout ensemble* E *dans un espace topologique* X *est fermée et est contenue dans chaque ensemble fermé de* X *qui contient* E, *de sorte que* $[E]$ *est le plus petit ensemble fermé de* X *qui contient* E. La première affirmation découle directement de 5° en vertu de 8° et, si $E \subset F \subset X$ et F est fermé, $[E] \subset [F] = F$ en vertu de 4° et de 8°.

10° *L'ensemble* E *d'un espace topologique est ouvert si et seulement si* $E = \,]E[$. En vertu de la définition 2, dire que E est ouvert signifie que E est un voisinage de chacun de ses points ou, ce qui revient au même de par la définition 5, que tous ses points sont intérieurs, i.e. $E \subset \,]E[$. Et en vertu de 3*, on a toujours $E \supset \,]E[$.

11° *L'ensemble E d'un espace topologique est fermé si et seulement si son complémentaire est ouvert.* En effet, en vertu des propositions 8°, 2° a) et 10°,

$$E \text{ est fermé} \Leftrightarrow E = [E] \Leftrightarrow \complement_X E = \complement_X [E] =]\complement_X E[\Leftrightarrow \complement_X E \text{ est ouvert.}$$

En remplaçant ici E par $\complement_X E$, nous obtenons la proposition duale suivante.

11*. *L'ensemble E d'un espace topologique est ouvert si et seulement si son complémentaire est fermé.*

EXERCICE 7. Montrer que l'intérieur de l'ensemble E dans un espace topologique est le plus grand ensemble ouvert contenu dans E.

6. Propriétés fondamentales des ensembles ouverts

Rappelons les opérations qu'on effectue sur les familles d'ensembles. Soit $(E_\alpha)_{\alpha \in A}$ une famille de sous-ensembles de l'ensemble X, i.e. une application de l'ensemble A des indices α dans l'ensemble de toutes les parties de l'ensemble X. La *réunion* et l'*intersection* des ensembles de cette famille sont respectivement les ensembles $\bigcup_{\alpha \in A} E_\alpha$ et $\bigcap_{\alpha \in A} E_\alpha$ définis de la façon suivante :

(1) $$x \in \bigcup_{\alpha \in A} E_\alpha \Leftrightarrow x \in E_\alpha \text{ pour un } \alpha \in A,$$

(2) $$x \in \bigcap_{\alpha \in A} E_\alpha \Leftrightarrow x \in E_\alpha \text{ pour tout } \alpha \in A.$$

Les réunions et les intersections sont liées entre elles par les formules

(3) $$\complement_X \bigcup_{\alpha \in A} E_\alpha = \bigcap_{\alpha \in A} \complement_X E_\alpha, \quad \complement_X \bigcap_{\alpha \in A} E_\alpha = \bigcup_{\alpha \in A} \complement_X E_\alpha.$$

En effet,

$$x \in \complement_X \bigcup_{\alpha \in A} E_\alpha \Leftrightarrow x \notin \bigcup_{\alpha \in A} E_\alpha \Leftrightarrow (\forall \alpha \in A)(x \notin E_\alpha) \Leftrightarrow$$
$$\Leftrightarrow (\forall \alpha \in A)(x \in \complement_X E_\alpha) \Leftrightarrow x \in \bigcap_{\alpha \in A} \complement_X E_\alpha,$$

et la deuxième formule (3) est obtenue de la première par substitution de $\complement_X E_\alpha$ à E_α et par passage aux complémentaires.

REMARQUE. Il se peut que l'ensemble A des indices d'une famille $(E_\alpha)_{\alpha \in A}$ de sous-ensembles de l'ensemble X soit vide. Quelles seront alors la réunion et l'intersection des ensembles de cette famille ? Leurs définitions (1) et (2) peuvent être notées comme suit :

(1') $$\bigcup_{\alpha \in \varnothing} E_\alpha = \{x \in X \mid \exists \alpha \in \varnothing, x \in E_\alpha\},$$

(2') $$\bigcap_{\alpha \in \varnothing} E_\alpha = \{x \in X \mid (\forall \alpha) \alpha \in \varnothing \Rightarrow x \in E_\alpha\}.$$

Comme la proposition ($\exists \alpha \in \varnothing, x \in E_\alpha$) dans (1′) est fausse, la proposition $x \in \bigcup_{\alpha \in \varnothing} E_\alpha$ est aussi fausse. Donc, $\bigcup_{\alpha \in \varnothing} E_\alpha = \varnothing$. D'autre part, l'implication $\alpha \in \varnothing \Rightarrow x \in E_\alpha$ dans (2′) étant vraie puisque sa prémisse est fausse, la proposition $x \in \bigcap_{\alpha \in \varnothing} E_\alpha$ est vraie pour tout $x \in X$. Donc,

$$\bigcap_{\alpha \in \varnothing} E_\alpha = X.$$

Comme nous le voyons, à la différence de la réunion, l'intersection d'une famille vide de sous-ensembles de l'ensemble X dépend sensiblement de X.

THÉORÈME 1. *La famille des ensembles ouverts d'un espace topologique X possède les propriétés suivantes* :

1) *la réunion de toute famille non vide d'ensembles ouverts est un ensemble ouvert* ;

2) *l'intersection de tout couple d'ensembles ouverts est un ensemble ouvert* ;

3) *l'ensemble vide est ouvert* ;

4) *tout le X est ouvert.*

Démonstration. 1) Soient $(O_\alpha)_{\alpha \in A}$ une famille non vide d'ensembles ouverts de l'espace X et $O = \bigcup_{\alpha \in A} O_\alpha$. Si $x \in O$, alors x appartient à un O_α ; étant ouvert, O_α est un voisinage du point x ; comme $O_\alpha \subset O$, O est aussi un voisinage du point x en vertu de l'axiome 1 des voisinages ; donc, O est un voisinage de chacun de ses points, i.e. un ensemble ouvert.

2) Soient O_1, O_2 des ensembles ouverts dans X et soit $O = O_1 \cap O_2$. Si $x \in O$, alors $x \in O_1$ et $x \in O_2$; par conséquent, O_1 et O_2 sont des voisinages du point x ; en vertu de l'axiome 2 des voisinages, O est aussi un voisinage de ce point ; donc, O est un voisinage de chacun de ses points, i.e. un ensemble ouvert.

3) L'implication « $x \in \varnothing \Rightarrow \varnothing$ est un voisinage du point x » est vraie puisque sa prémisse « $x \in \varnothing$ » est fausse. Autrement dit, l'ensemble vide est un voisinage de chacun de ses points (faute de ces derniers), i.e. un ensemble ouvert.

4) Soit $x \in X$. En vertu de la définition 1, l'ensemble des voisinages du point x n'est pas vide. Soit U l'un de ces voisinages. Comme $U \subset X$, X est aussi un voisinage de x en vertu de l'axiome 1 des voisinages. Vu qu'il est voisinage de chacun de ses points, X est un ensemble ouvert.

REMARQUE. $\varnothing$ est l'unique ensemble ouvert dans l'espace topologique $\varnothing$.

Les propriétés 1 à 4 caractérisent bien l'ensemble des ouverts de l'espace topologique et la topologie elle-même. Plus précisément, on a le théorème suivant.

THÉORÈME 2. *Soit $\mathcal{T}$ un ensemble de parties de l'ensemble X possédant les propriétés suivantes* :

1) *la réunion de toute famille non vide d'ensembles de $\mathcal{T}$ appartient à $\mathcal{T}$;*

2) *l'intersection de deux ensembles quelconques de $\mathcal{T}$ appartient à $\mathcal{T}$;*

3) $\varnothing \in \mathcal{T}$;

4) $X \in \mathcal{T}$.

Il s'ensuit que $\mathcal{T}$ est un ensemble des ouverts pour la topologie univoquement définie sur X. Cette topologie est donnée par les voisinages des points, qui sont définis de la façon suivante : U est appelé voisinage du point a s'il existe un $O \in \mathcal{T}$ tel que $a \in O \subset U$.

Démonstration. Pour $X = \varnothing$ le théorème est évident. Soit $X \neq \varnothing$. Notons $\mathcal{U}(a)$ l'ensemble des voisinages ainsi définis du point $a \in X$. Cet ensemble n'est pas vide. En effet, $X \in \mathcal{T}$ d'après la propriété 4, et comme $a \in X \subset X$, on a $X \in \mathcal{U}(a)$. Vérifions les axiomes des voisinages.

1) Soit $U \subset V \subset X$. Si $U \in \mathcal{U}(a)$, il existe par définition un $O \in \mathcal{T}$ tel que $a \in O \subset U$. Donc, $a \in O \subset V$ et, par conséquent, $V \in \mathcal{U}(a)$.

2) Soient $U_1 \in \mathcal{U}(a)$ et $U_2 \in \mathcal{U}(a)$, de sorte qu'il existe des ensembles $O_1 \in \mathcal{T}$ et $O_2 \in \mathcal{T}$ tels que $a \in O_1 \subset U_1$ et $a \in O_2 \subset U_2$. Il s'ensuit que $a \in O_1 \cap O_2 \subset U_1 \cap U_2$ et comme $O_1 \cap O_2 \in \mathcal{T}$ en vertu de la propriété 2, on a $U_1 \cap U_2 \in \mathcal{U}(a)$.

3), 4) Soit $U \in \mathcal{U}(a)$. Il existe un $O \in \mathcal{T}$ tel que $a \in O \subset U$, d'où $a \in U$. Vu que $O \subset O$, O est par définition un voisinage de tout point $x \in O$.

Ainsi, tous les axiomes des voisinages sont vérifiés et chaque $O \in \mathcal{T}$ est un ensemble ouvert pour la topologie t donnée par ce système de voisinages. Soit maintenant V un ensemble arbitraire non vide ouvert pour cette topologie. Comme V est un voisinage de chacun de ses points, pour tout $x \in V$ l'ensemble des $O \in \mathcal{T}$ tels que $x \in O \subset V$ n'est pas vide. Soit O_x la réunion de ces O. La propriété 1 dit que $O_x \in \mathcal{T}$ et, en vertu de la définition de l'ensemble O_x, $x \in O_x \subset V$. Il s'ensuit que $V = \bigcup_{x \in V} O_x$ et, par conséquent, $V \in \mathcal{T}$ en vertu de la propriété 1. Ainsi donc, $\mathcal{T}$ coïncide avec l'ensemble des ouverts pour la topologie t. Supposons enfin que $\mathcal{T}$ soit un ensemble des ouverts pour une autre topologie t'. Si U' est un voisinage du point a pour cette topologie, l'axiome 4 des voisinages dit que U' contient un voisinage ouvert du point a, i.e. il existe un ensemble $O \in \mathcal{T}$ tel que $a \in O \subset U'$, ce qui implique $U' \in \mathcal{U}(a)$. Réciproquement : si $U \in \mathcal{U}(a)$, U est un voisinage du point a pour la topologie t' car il contient un voisinage ouvert O du point a pour cette topologie. Ainsi, les ensembles des voisinages des points pour les topologies t et t' coïncident et, par conséquent, $t' = t$ en vertu de la définition 1.

On prend le plus souvent pour la notion topologique préliminaire non pas un voisinage mais un ensemble ouvert et on définit l'espace topologique comme un ensemble dans lequel est donné l'ensemble $\mathcal{T}$ de parties appelées ensembles ouverts, qui vérifie les conditions du théorème 2. Cependant, vu que nous aurons affaire surtout aux espaces métriques, il serait plus commode de définir la topologie en recourant à la notion de voisinage.

Les ensembles fermés et ouverts étant complémentaires les uns des autres (propositions 11° et 11*), le théorème 1 admet le théorème dual suivant.

THÉORÈME 1*. *L'ensemble des sous-ensembles fermés d'un espace topologique X possède les propriétés suivantes :*

1) *l'intersection de toute famille non vide d'ensembles fermés est un ensemble fermé ;*

2) *la réunion de deux ensembles fermés quelconques est un ensemble fermé ;*

3) *tout l'espace X est fermé ;*

4) *l'ensemble vide est fermé.*

Démonstration. 1) Soit $(F_\alpha)_{\alpha\in A}$ une famille non vide d'ensembles fermés. Comme les ensembles $\complement_X F_\alpha$ sont ouverts, la proposition 1) du théorème 1 entraîne que $\bigcup_{\alpha\in A} \complement_X F_\alpha$ est aussi ouvert. Mais alors $\bigcap_{\alpha\in A} F_\alpha = \complement_X\left(\bigcup_{\alpha\in A} \complement_X F_\alpha\right)$ est un ensemble fermé.

2) Soient F_1 et F_2 des ensembles fermés. Il s'ensuit que $\complement_X F_1$ et $\complement_X F_2$ sont ouverts et, suivant la proposition 2) du théorème 1, $\complement_X F_1 \cap \complement_X F_2$ est aussi ouvert. Par conséquent, $F_1 \cup F_2 = \complement_X(\complement_X F_1 \cap \complement_X F_2)$ est fermé.

3) Comme $\varnothing$ est ouvert d'après la proposition 3) du théorème 1, $X = \complement_X \varnothing$ est fermé.

4) Comme X est ouvert d'après la proposition 4) du théorème 1, $\varnothing = \complement_X X$ est fermé.

REMARQUE. $\varnothing$ est l'unique ensemble fermé de l'espace topologique $\varnothing$.

EXERCICE 8. Formuler et démontrer le théorème 2* dual au théorème 2, qui montre que la topologie est aussi définie par une famille d'ensembles fermés. Montrer que cette dernière peut être donnée, à son tour, par l'opération de passage à l'ensemble fermé. Plus précisément, soit donnée sur l'ensemble $\mathscr{P}(X)$ des parties de l'ensemble X l'opération $E \to [E]$ possédant les propriétés 3° à 6°. Appelons ensemble fermé dans X tout $E \in \mathscr{P}(X)$ (i.e. $E \subset X$) coïncidant avec $[E]$. Montrer que la famille de ces ensembles coïncide avec celle des ensembles fermés pour la topologie définie univoquement sur X et que $E \to [E]$ l'est avec l'opération de passage à un ensemble fermé pour cette topologie.

7. Sous-espace

Soient X un ensemble et E son sous-ensemble. L'intersection de tout ensemble $A \subset X$ avec E s'appelle *trace* de A sur E.

THÉORÈME 3. *Soient X un espace topologique et $E \subset X$. Si les traces sur E de tous les voisinages dans X de chaque point $a \in E$ sont les voisinages du point a dans E, le système ainsi défini des voisinages des points dans E vérifie les axiomes des voisinages et transforme E en espace topologique.*

Démonstration. Désignons par $\mathscr{U}(a)$ l'ensemble des voisinages du point $a \in E$ dans X et par $\mathscr{U}_E(a)$ celui dans E. Comme $X \in \mathscr{U}(a)$ (la fin de la démonstration du théorème 1), on a $E = X \cap E \in \mathscr{U}_E(a)$, de sorte que $\mathscr{U}_E(a)$ n'est pas vide. Vérifions les axiomes des voisinages.

1) Soit $V \subset W \subset E$. Si $V \in \mathscr{U}_E(a)$, alors $V = U \cap E$ où $U \in \mathscr{U}(a)$. Or

$$W = V \cup W = (U \cap E) \cup (W \cap E) = (U \cup W) \cap E.$$

Il s'ensuit que $W \in \mathscr{U}_E(a)$ car $U \cup W \in \mathscr{U}(a)$ (axiome 1 des voisinages).

2) Soient $V_1, V_2 \in \mathscr{U}_E(a)$, de sorte que $V_1 = U_1 \cap E$ et $V_2 = U_2 \cap E$ où $U_1, U_2 \in \mathscr{U}(a)$. Alors $V_1 \cap V_2 = (U_1 \cap U_2) \cap E$. Or $U_1 \cap U_2 \in \mathscr{U}(a)$ (axiome 2 des voisinages). Par conséquent, $V_1 \cap V_2 \in \mathscr{U}_E(a)$.

3) Soit $V \in \mathscr{U}_E(a)$, de sorte que $V = U \cap E$ où $U \in \mathscr{U}(a)$. On a $a \in U$ en vertu de l'axiome 3 des voisinages, et $a \in E$ par hypothèse. Donc, $a \in V$.

4) Soit $V \in \mathscr{U}_E(a)$. Il s'ensuit comme dans le cas précédent que $V = U \cap E$ où $U \in \mathscr{U}(a)$. En vertu de l'axiome 4 des voisinages, U contient un voisinage ouvert O du point a dans X, de sorte que $V \supset O \cap E$. Mais comme O est un voisinage de chacun de ses points dans X, il en est de même pour $O \cap E$ dans E.

DÉFINITION 7. Soient X un espace topologique et t sa topologie. On dit que le sous-ensemble $E \subset X$ muni de la topologie décrite dans le théorème 3 est un *sous-espace topologique* (ou *sous-espace* tout court) *de l'espace* X et que cette topologie sur E est une *topologie induite* par celle de X ; elle sera notée $t\,|_E$.

EXERCICE 9. Vérifier que $\mathbf{R}$ est un sous-espace de l'espace $\bar{\mathbf{R}}$, i.e. que sa topologie coïncide avec la topologie induite par celle de $\bar{\mathbf{R}}$.

EXERCICE 10. Vérifier que $\mathbf{R}^n$ est un sous-espace de l'espace $\widehat{\mathbf{R}^n}$.

EXERCICE 11. $\mathbf{R}$ peut être considéré comme partie de $\mathbf{R}^2$ si l'on identifie, par exemple, les points $x \in \mathbf{R}$ avec les couples $(x\,;\,0) \in \mathbf{R}^2$. Vérifier que par cette identification $\mathbf{R}$ devient un sous-espace de l'espace $\mathbf{R}^2$.

THÉORÈME 4. *Si E est un sous-espace de l'espace topologique X, l'adhérence $[A]_E$ de tout ensemble $A \subset E$ dans E est la trace sur E de l'adhérence $[A]$ de cet ensemble dans X, i.e. $[A]_E = [A] \cap E$. Autrement dit, x est un point adhérent à A dans E si et seulement si x est un point adhérent à A dans X et si x appartient à E.*

Démonstration. $x \in [A]_E$ signifie que $x \in E$ et que chaque voisinage U du point x dans E a une intersection non vide avec A. Mais U est un voisinage du point x dans E si et seulement si $U = V \cap E$ où V est un voisinage du point x dans X. Par conséquent, $x \in [A]_E$ si et seulement si $x \in E$ et si $V \cap A = V \cap E \cap A \neq \varnothing$ pour chaque voisinage V du point x dans X, i.e. $x \in [A] \cap E$.

REMARQUE. Une proposition analogue pour l'intérieur de A n'est pas juste. Par exemple, si $X = \mathbf{R}$ et E est un intervalle $[a, b]$, l'intérieur $]A[_E$ de l'ensemble $A = E$ dans E est le segment $[a, b]$, tandis que l'intérieur de A dans X est l'intervalle $]a, b[$, de sorte que $]A[_E \neq\,]A[\cap E$.

THÉORÈME 5. *Soit E un sous-espace de l'espace topologique X et soit $A \subset E$. Alors :*

1) *A est fermé dans E si et seulement si A est la trace sur E d'un ensemble fermé $F \subset X$; ceci étant, $[A]$ est le plus petit des ensembles F ;*

2) *A est ouvert dans E si et seulement si A est la trace sur E d'un ensemble ouvert $O \subset X$; ceci étant, $X \setminus [E \setminus A]$ est le plus grand des ensembles O.*

Démonstration. 1) Si A est fermé dans E, i.e. $A = [A]_E$, alors $A = [A] \cap E$ en vertu du théorème 4, si bien que A est la trace sur E de

voisinage U du point a (de $\mathscr{U}$) tel que $f(x) \in V'$ pour tout $x \in U \cap D_f$; donc, à plus forte raison, $f(x) \in V$ quel que soit $x \in U \cap D_f$, i.e. $f(x) \to b$ quand $x \to a$.

EXEMPLE 1. En vertu de la définition 1.2.2, la suite (x_n) de points d'un espace métrique muni de la métrique d converge vers un point a de cet espace si (et seulement si) $d(x_n, a) \to 0$. Envisageons maintenant cette suite comme application de $\bar{\mathbf{R}}$ dans X, avec $D_{(x_n)} = \mathbf{N}$. On peut parler de sa limite quand $n \to +\infty$ car $+\infty \in \bar{\mathbf{N}}$ (si bien que la condition 1) de la définition 1 et du théorème 1 est vérifiée). Prenons pour le système fondamental $\mathscr{U}$ de voisinages du point $+\infty$ l'ensemble des intervalles $]M, +\infty]$ $(M > 0)$ et pour le système fondamental $\mathscr{V}$ de voisinages du point a l'ensemble de ses ε-voisinages $U(a\,;\varepsilon)$ $(\varepsilon > 0)$. En vertu du théorème 1, $x_n \to a$ signifie que

$$(\forall \varepsilon > 0)(\exists M > 0)(\forall n \in]M, +\infty] \cap \mathbf{N})(x_n \in U(a\,;\varepsilon)),$$

i.e. pour tout $\varepsilon > 0$ il existe un $M > 0$ tel que $d(x_n, a) < \varepsilon$ pour tous les entiers naturels $n > M$. Or cela signifie justement que $d(x_n, a) \to 0$. Donc, $x_n \to a$ si et seulement si la suite (x_n) converge vers a.

EXEMPLE 2. Soient X et Y des espaces métriques munis des métriques d_X et d_Y respectivement, f une application de X dans Y, $a \in X$, $b \in Y$. Prenons pour $\mathscr{U}$ et $\mathscr{V}$ les ensembles des voisinages sphériques des points a et b respectivement. Nous obtenons que $f(x) \to b$ quand $x \to a$ si et seulement si $a \in \overline{D_f}$ et

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in U(a\,;\delta) \cap D_f)(f(x) \in V(b\,;\varepsilon)),$$

i.e. $(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in D_f)(d_X(x, a) < \delta \Rightarrow d_Y(f(x), b) < \varepsilon)$.

Nous verrons par la suite qu'il est plus commode de formuler la définition 1 de la limite en termes de la théorie des ensembles. Pour le faire, il faut posséder une technique permettant d'appliquer la notion de fonction aux ensembles.

2. Image directe et image réciproque d'un ensemble.

Toute application f d'un ensemble X dans un ensemble Y engendre une application de l'ensemble $\mathscr{P}(X)$ des parties de X dans l'ensemble $\mathscr{P}(Y)$ des parties de Y, qui est définie pour tout $E \in \mathscr{P}(X)$ (i.e. $E \subset X$) par la formule

$$f(E) = \{f(x) \mid x \in E \cap D_f\}.$$

Il s'ensuit que $f(E) = f(E \cap D_f)$, d'où $f(E) = \varnothing \Leftrightarrow E \cap D_f = \varnothing$ et, en particulier, $f(\varnothing) = \varnothing$. On dit que $f(E)$ est l'*image de l'ensemble E par l'application f*; en particulier, $R_f = f(x) = f(D_f)$. On montre aisément que

$$E \subset E' \Rightarrow f(E) \subset f(E'), \quad f(E_1 \cup E_2) = f(E_1) \cup f(E_2),$$
$$f(E_1 \cap E_2) \subset f(E_1) \cap f(E_2).$$

Les deux dernières propriétés s'étendent respectivement à la réunion et à l'intersection de n'importe quelle famille d'ensembles :

1) $f\left(\bigcup_{\alpha\in A} E_\alpha\right) = \bigcup_{\alpha\in A} f(E_\alpha)$. En effet, $y\in f\left(\bigcup_{\alpha\in A} E_\alpha\right) \Leftrightarrow y = f(x)$ où $x\in\left(\bigcup_{\alpha\in A} E_\alpha\right)\cap D_f = \bigcup_{\alpha\in A}(E_\alpha\cap D_f) \Leftrightarrow y = f(x)$ avec $x\in E_\alpha\cap D_f$ pour un $\alpha\in A \Leftrightarrow y\in f(E_\alpha)$ pour $\alpha\in A \Leftrightarrow y\in \bigcup_{\alpha\in A} f(E_\alpha)$;

2) $f\left(\bigcap_{\alpha\in A} E_\alpha\right) \subset \bigcap_{\alpha\in A} f(E_\alpha)$. En effet, soit $E = \bigcap_{\alpha\in A} E_\alpha$. Comme $E\subset E_\alpha$ pour tout $\alpha\in A$, il vient $f(E)\subset f(E_\alpha)$ pour tout $\alpha\in A$, i.e. $f(E)\subset \bigcap_{\alpha\in A}(E_\alpha)$.

Pour tout $F\in\mathscr{P}(Y)$ (i.e. $F\subset Y$) posons

$$f^{-1}(F) = \{x\in X \mid x\in D_f \text{ et } f(x)\in F\}$$

(de sorte que $f^{-1}(F)\in\mathscr{P}(X)$ et f^{-1} est une application de $\mathscr{P}(Y)$ dans $\mathscr{P}(X)$). On dit que $f^{-1}(F)$ est l'*image réciproque de l'ensemble F par l'application f*. Ainsi, $D_f = f^{-1}(Y) = f^{-1}(R_f)$. Il est évident que $f^{-1}(F) = f^{-1}(F\cap R_f)$, donc $f^{-1}(F) = \varnothing \Leftrightarrow F\cap R_f = \varnothing$ et, en particulier, $f^{-1}(\varnothing) = \varnothing$.

Notons quelques propriétés de l'application $f^{-1}: \mathscr{P}(Y)\to\mathscr{P}(X)$.

1) $F\subset F' \Rightarrow f^{-1}(F)\subset f^{-1}(F')$. En effet, si $x\in f^{-1}(F)$, i.e. $x\in D_f$ et $f(x)\in F$, alors, à plus forte raison, $f(x)\in F'$, i.e. $x\in f^{-1}(F')$.

2) $f^{-1}\left(\bigcup_{\alpha\in A} F_\alpha\right) = \bigcup_{\alpha\in A} f^{-1}(F_\alpha)$. En effet, $x\in f^{-1}\left(\bigcup_{\alpha\in A} F_\alpha\right) \Leftrightarrow x\in D_f$ et $f(x)\in \bigcup_{\alpha\in A} F_\alpha \Leftrightarrow x\in D_f$ et $f(x)\in F_\alpha$ pour un $\alpha \Leftrightarrow x\in f^{-1}(F_\alpha)$ pour un $\alpha \Leftrightarrow x\in \bigcup_{\alpha\in A} f^{-1}(F_\alpha)$.

3) $f^{-1}\left(\bigcap_{\alpha\in A} F_\alpha\right) = \bigcap_{\alpha\in A} f^{-1}(F_\alpha)$ (à la différence des images, on a ici $=$ et non seulement $\subset$!). En effet, $x\in f^{-1}\left(\bigcap_{\alpha\in A} F_\alpha\right) \Leftrightarrow x\in D_f$ et $f(x)\in \bigcap_{\alpha\in A} F_\alpha \Leftrightarrow x\in D_f$ et $f(x)\in F_\alpha$ pour tout $\alpha \Leftrightarrow x\in f^{-1}(F_\alpha)$ pour tout $\alpha \Leftrightarrow x\in \bigcap_{\alpha\in A} f^{-1}(F_\alpha)$.

de la limite de la somme, du produit, du quotient et leurs complémentaires ; les théorèmes de passage à la limite dans les inégalités.

Les théorèmes de la limite de la somme de deux fonctions et du produit de la fonction par un nombre sont aussi valables pour les fonctions à valeurs dans un espace normé. Soient X un espace topologique, Y un espace vectoriel réel et f, g des applications de X dans Y. On entend par $f+g$ une application de X dans Y qui se définit de la façon suivante : $D_{f+g} = D_f \cap D_g$ et $(f+g)(x) = f(x) + g(x)$ pour tout $x \in D_{f+g}$. On définit de façon analogue λf (avec $D_{\lambda f} = D_f$) où $f: X \to Y$ et $\lambda \in \mathbf{R}$.

THÉORÈME 5. *Soient X un espace topologique, Y un espace normé, f, g des applications de X dans Y, $a \in X$ et $b, c \in Y$. Si $f(x) \to b$ quand $x \to a$, $g(x) \to c$ quand $x \to a$ et $a \in \overline{D_{f+g}}$, alors $(f+g)(x) \to b + c$ quand $x \to a$.*

Démonstration. Pour tout $\varepsilon > 0$ il existe des voisinages U' et U'' du point a tels que

$$x \in D_f \cap U' \Rightarrow \|f(x) - b\| < \frac{\varepsilon}{2},$$

$$x \in D_g \cap U'' \Rightarrow \|g(x) - c\| < \frac{\varepsilon}{2}.$$

Soit $U = U' \cap U''$. L'ensemble U est aussi un voisinage du point a et si $x \in D_{f+g} \cap U$ $(= (D_f \cap U') \cap (D_g \cap U''))$, on obtient

$$\|(f+g)(x) - (b+c)\| \leqslant \|f(x) - b\| + \|g(x) - c\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

EXERCICE 2. Montrer que, pour tout $\lambda \in \mathbf{R}$, $\lambda f(x) \to \lambda b$ quand $x \to a$ si $f(x) \to b$ quand $x \to a$.

Notons en conclusion que le **théorème de la limite d'une application constante** est aussi vrai : *si X et Y sont des espaces topologiques, f l'application de X dans Y à une valeur constante c et $a \in \overline{D_f}$, alors $f(x) \to c$ quand $x \to a$.*

§ 2. Applications continues

1. Notion d'application continue

L'application continue est définie tout comme les applications de $\mathbf{R}$ dans $\mathbf{R}$.

DÉFINITION 1. Soient X et Y des espaces topologiques. Une application f de X dans Y est dite *continue en un point* x_0 si $x_0 \in D_f$ et $f(x) \to f(x_0)$ quand $x \to x_0$. L'application est dite *continue* si elle est continue en tout point $x_0 \in D_f$.

Ainsi donc, en vertu du théorème 2 du § 1, *f est continue en un point x_0 si et seulement si $x_0 \in D_f$ et $f^{-1}(V)$ est un voisinage du point x_0 dans D_f pour tout voisinage V du point $f(x_0)$ dans Y.*

Citons deux exemples élémentaires d'applications continues.

EXEMPLE 1. *L'application constante est continue* : si $f(x) = c$ pour tout $x \in D_f$, alors $f(x) \to c$ quand $x \to x_0$ quel que soit le point $x_0 \in D_f$.

EXEMPLE 2. *L'injection canonique d'un sous-espace E de l'espace X dans X est continue.* Soit i une injection canonique de E dans X, i.e. une application de E dans X qui fait correspondre à chaque point $x \in E$ le même point envisagé comme élément de X. Pour tout point $x \in E$ et tout voisinage U du point $i(x) = x$ dans X on a alors

$$i^{-1}(U) = U \cap E,$$

de sorte que $i^{-1}(U)$ est un voisinage du point x dans E. Par conséquent, i est continue en chaque point $x \in E$.

Si X et Y sont des espaces métriques et d_X et d_Y leurs métriques respectives, l'application f de X dans Y est continue au point x_0 si et seulement si $x_0 \in D_f$ et

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in D_f)(d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon)$$

(voir exemple 2 du § 1).

EXEMPLE 3. Soient X un espace normé réel (définition 1.2.3) et a son point. L'application $f : \mathbf{R} \to X$ définie par la formule $f(\lambda) = \lambda a$ est continue. En effet, pour $a = 0$ elle est constante et pour $a \neq 0$ (i.e. $\|a\| \neq 0$) et tout $\lambda_0 \in \mathbf{R}$ on a $\|\lambda a - \lambda_0 a\| = |\lambda - \lambda_0| \, \|a\| < \varepsilon$ si $|\lambda - \lambda_0| < \dfrac{\varepsilon}{\|a\|}$.

2. Propriétés fondamentales des applications continues

THÉORÈME 1 (passage à la limite sous le symbole de la fonction continue). *Soient X, Y, Z des espaces topologiques, f une application de X dans Y, g une application de Y dans Z, $a \in X$ et $y_0 \in Y$. Si $f(x) \to y_0$ quand $x \to a$, g est continue au point y_0 et $a \in D_{g \circ f}$, on a $g(f(x)) \to g(y_0)$ quand $x \to a$.*

Démonstration. La continuité de g en y_0 signifie que $y_0 \in D_g$ et $g(y) \to g(y_0)$ quand $y \to y_0$. Il reste à appliquer le théorème de la limite de la composée des applications.

COROLLAIRE. *Les notations étant celles du théorème* 1, *si g est continue au point y_0 et (y_n) est une suite de points dans D_g qui converge vers y_0, on a $g(y_n) \to g(y_0)$.*

Démonstration. Posons $X = \overline{\mathbf{R}}$, $D_f = \mathbf{N}$, $f(n) = y_n$, $a = +\infty$ et appliquons le théorème 1.

THÉORÈME 2 (composée d'applications continues). *Soient X, Y, Z des espaces topologiques, f une application de X dans Y et g une application de Y dans Z. Si f est continue en un point x_0, et g au point $y_0 = f(x_0)$, $g \circ f$ est continue en x_0.*

Démonstration. Posons $a = x_0$; la continuité de f en x_0 signifie alors que $f(x) \to y_0$ quand $x \to a$. En outre, $a \in D_{g \circ f}$, donc, à plus forte raison,

i.e. $c \in U(x_0 ; r)$ pour tout $c \in [a, b]$. La convexité de $B(x_0 ; r)$ est démontrée de façon analogue si on remplace le signe $<$ par $\leqslant$ dans la dernière inégalité.

2. Ensemble et application connexes

Les notions d'ensemble connexe dans $\mathbf{R}$ et de fonction connexe de $\mathbf{R}$ dans $\mathbf{R}$ peuvent être facilement généralisées aux sous-ensembles et applications de tous espaces topologiques.

DÉFINITION 3. Un ensemble E de l'espace topologique X est dit *connexe* si dans toute partition de E en deux sous-ensembles non vides, l'un d'eux au moins contient un point adhérent à l'autre.

On montre que dans $\mathbf{R}$ seuls les intervalles sont connexes. Mais déjà dans $\mathbf{R}^2$ la collection des ensembles connexes est beaucoup plus riche.

DÉFINITION 4. Soient X et Y des espaces topologiques. Une application f de X dans Y est dite *connexe* si $f(E)$ est connexe dans Y pour tout ensemble connexe $E \subset D_f$.

THÉORÈME 1. *Toute application continue est connexe.*

DÉFINITION 5. On appelle *courbe de Jordan dans* $\mathbf{R}^n$ l'image par toute application continue d'un intervalle non dégénéré I dans $\mathbf{R}^n$. L'image par restriction de cette application à un segment arbitraire $\Delta \subset I$ est appelée *arc de Jordan.*

Comme les intervalles sont des ensembles connexes, le théorème 1 entraîne le corollaire suivant.

COROLLAIRE. *L'ensemble des points de la courbe de Jordan est connexe.*

En particulier, *tout segment* $[a, b]$ *dans* $\mathbf{R}^n$ *est connexe*. En effet, l'application $\varphi : [0, 1] \to \mathbf{R}^n$ qui donne, dans la définition 1, le segment $[a, b]$ est continue en tant que somme des restrictions à $[0, 1]$ des applications continues $\lambda \mapsto (1 - \lambda)a$ et $\lambda \mapsto \lambda b$ de $\mathbf{R}$ dans $\mathbf{R}^n$ (voir exemple 2.2.3).

THÉORÈME 2 (théorème des valeurs intermédiaires). *Soit f une fonction réelle continue sur un sous-ensemble connexe E de l'espace topologique. Si f prend des valeurs distinctes aux points $a, b \in E$, tout nombre C compris entre $f(a)$ et $f(b)$ est valeur de f en un point au moins $c \in E$.*

Démonstration. D'après le théorème 1, $f(E)$ est un ensemble connexe dans $\mathbf{R}$, i.e. est un intervalle et $f(a)$ et $f(b)$ sont ses points. Comme l'intervalle est un ensemble convexe, $C \in f(E)$, i.e. $C = f(c)$ pour un $c \in E$.

DÉFINITION 6. Un espace topologique X est dit *connexe* s'il est connexe en tant qu'ensemble dans X.

THÉORÈME 3. *L'espace topologique est connexe si et seulement s'il n'existe pas de partition de cet espace en deux ouverts non vides.*

Démonstration. Supposons qu'il existe une partition de X en deux ouverts O_1 et O_2. Si $x \in O_1$, on a $x \notin [O_2]$ parce que x possède un voisinage O_1 qui ne se coupe pas avec O_2. De façon analogue, $x \notin [O_1]$ si $x \in O_2$. Ainsi

donc, aucun des ensembles O_1, O_2 ne contient le point adhérent à l'autre. Pour cette raison, si X est connexe, on a $O_1 = \varnothing$ ou $O_2 = \varnothing$, de sorte qu'il n'existe aucune partition de X en deux ouverts non vides. Réciproquement : supposons que X possède cette propriété et $X = A_1 \cup A_2$ où $A_1 \cap A_2 = \varnothing$. Si $A_1 \cap [A_2] = \varnothing$, on obtient $[A_2] \subset A_2$, i.e. A_2 est fermé et, par conséquent, A_1 est ouvert. Si en outre $[A_1] \cap A_2 = \varnothing$, on obtient de façon analogue que A_2 est ouvert. Mais alors $A_1 = \varnothing$ ou $A_2 = \varnothing$. Ainsi dans n'importe quelle partition de l'espace X en deux sous-ensembles non vides, l'un d'eux au moins doit contenir un point adhérent à l'autre, i.e. X est connexe.

EXERCICE 2. Montrer que l'espace topologique est connexe si et seulement s'il n'existe aucune partition de cet espace en deux ensembles fermés non vides.

EXERCICE 3. Un sous-ensemble de l'espace topologique X est dit *ouvert-fermé* s'il est à la fois fermé et ouvert. Tels sont toujours $\varnothing$ et tout le X. Montrer que X est connexe si et seulement si $\varnothing$ et X sont les seuls ensembles ouverts-fermés dans X.

EXERCICE 4. Montrer qu'un sous-ensemble de l'espace topologique est connexe si et seulement s'il est connexe comme sous-espace.

EXERCICE 5. Montrer que l'adhérence de l'ensemble connexe est connexe. Est-ce que la proposition analogue pour l'intérieur est vraie ?

THÉORÈME 4. *Soit E un sous-espace de l'espace topologique X. Un ensemble $F \subset E$ est connexe dans E si et seulement s'il est connexe dans X.*

Démonstration. Soit $F = A \cup B$ où $A \neq \varnothing$, $B \neq \varnothing$, $A \cap B = \varnothing$. En vertu du théorème 1.4.4,

$$A \cap [B]_E = A \cap [B] \cap E = A \cap [B]$$

et $[A]_E \cap B = [A] \cap B$. Pour cette raison, F étant connexe dans E, on obtient $A \cap [B] \neq \varnothing$ ou $[A] \cap B \neq \varnothing$ pour toute partition de l'ensemble F, i.e. F est connexe dans X.

THÉORÈME 5. *Pour qu'une partie E de l'espace topologique X soit connexe il est nécessaire et suffisant que deux points quelconques a et b de E soient joints par un ensemble connexe contenu dans E, i.e. qu'il existe dans X un ensemble connexe F tel que $a \in F$, $b \in F$ et $F \subset E$.*

Démonstration. Il est évident que cette condition est nécessaire : si E est connexe, pour F on peut prendre E lui-même. Démontrons qu'elle est suffisante. Supposons que E vérifie les hypothèses du théorème et que $E = A \cup B$ où $A \neq \varnothing$, $B \neq \varnothing$, $A \cap B = \varnothing$. Choisissons des points $a \in A$, $b \in B$ et considérons un ensemble F connexe dans X tel que $F \subset E$ et $a \in F$, $b \in F$. Supposons que $A_1 = A \cap F$, $B_1 = B \cap F$. Alors : 1) $A_1 \cup B_1 = (A \cup B) \cap F = E \cap F = F$; 2) $a \in A_1$, $b \in B_1$, de sorte que $A_1 \neq \varnothing$, $B_1 \neq \varnothing$; 3) $A_1 \cap B_1 \subset A \cap B$, de sorte que $A_1 \cap B_1 = \varnothing$. Comme F est connexe, on a $A_1 \cap [B_1] \neq \varnothing$ ou $[A_1] \cap B_1 \neq \varnothing$, et comme $A_1 \subset A$ et $B_1 \subset B$, on obtient à plus forte raison que $A \cap [B] \neq \varnothing$ ou $[A] \cap B \neq \varnothing$, donc E est connexe.

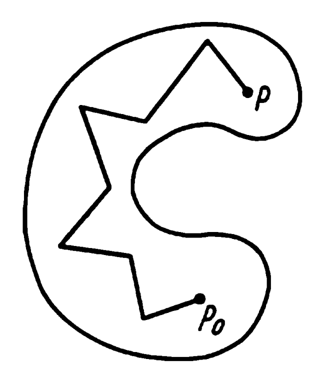


Fig. 7

la boule $U(p_0\,;\,\delta_0)$ est convexe (voir n° 1), le segment $[p_0,\,p]$ est contenu dans $U(p_0\,;\,\delta_0)$, donc dans D pour tout point $p \in U(p_0\,;\,\delta_0)$, i.e. p est joint à p_0 ; donc $U(p_0\,;\,\delta_0) \subset D_0$. Par ailleurs, D_0 est un ensemble ouvert. En effet, soit $p \in D_0$, i.e. p est joint à p_0. Puisque D est ouvert, il existe un $\delta > 0$ tel que $U(p\,;\,\delta) \subset D$. Or tout point $p' \in U(p\,;\,\delta)$ est joint à p par un segment contenu dans $U(p\,;\,\delta)$, donc dans D. En joignant ce segment à la ligne polygonale dans D qui joint p à p_0, on obtient une ligne polygonale dans D qui joint p' à p_0 ; ainsi donc, $U(p\,;\,\delta) \subset D_0$. Posons $D_1 = D \setminus D_0$; D_1 est aussi un ensemble ouvert. En effet, soit $p_1 \in D_1$; il existe un $\delta_1 > 0$ tel que $U(p_1\,;\,\delta_1) \subset D$; p_1 est joint à tout point $p \in U(p_1\,;\,\delta_1)$. C'est pourquoi, si $p \in D_0$, i.e. p était joint à p_0, le point p_1 serait joint à p_0, tandis que $p_1 \in D_0$; donc, $U(p_1\,;\,\delta_1) \cap D_0 = \varnothing$, i.e. $U(p_1\,;\,\delta_1) \subset D_1$, de sorte que D_1 est ouvert. Soit maintenant D un domaine. On a $D = D_0$. En effet, autrement on aurait une partition de D en deux ouverts non vides D_0 et D_1, ce qui contredit la connexité de D parce que chaque point de D_0 serait séparé de D_1 par son voisinage D_0 et, par conséquent, ne pourrait pas être le point adhérent à D_1 ; pour la même raison, aucun point de D_1 ne pourrait être un point adhérent à D_0. Donc, n'importe quels deux points $p,\,p' \in D$ sont joints à p_0, de sorte qu'ils sont joints l'un à l'autre par une ligne brisée contenue dans D, i.e. D est connexe par lignes brisées.

THÉORÈME 10. *Tout ouvert non vide de* $\mathbf{R}^n$ *est une réunion de ses domaines disjoints deux à deux.*

Démonstration. Soit O un ouvert non vide de $\mathbf{R}^n$. Considéré comme sous-espace de $\mathbf{R}^n$, l'ensemble O admet en vertu du théorème 7 une partition en composantes connexes. Soit G l'une d'elles et soit $x \in G$. Etant ouvert dans $\mathbf{R}^n$, O contient un voisinage sphérique U du point x. En tant qu'ensemble connexe dans $\mathbf{R}^n$ (voir n° 2), U est aussi connexe dans O (théorème 4). Par conséquent, $U \subset G$ puisque G est la réunion de tous les ensembles connexes dans O contenant x. De ce fait, G est un ensemble ouvert dans $\mathbf{R}^n$. D'après le même théorème 4, G est aussi connexe dans $\mathbf{R}^n$. Ainsi

donc, les composantes connexes de O sont des ensembles ouverts connexes dans $\mathbf{R}^n$, i.e. les domaines, et O est une réunion de ces domaines (disjoints deux à deux).

COROLLAIRE. *Tout ouvert non vide de* $\mathbf{R}$ *est une réunion de ses intervalles ouverts disjoints deux à deux.*

En effet, les domaines dans $\mathbf{R}$ sont des intervalles ouverts.

§ 2. Compacité *)

1. Théorème de Borel-Lebesgue

En étudiant la théorie de la mesure et le calcul intégral, les mathématiciens français Borel et Lebesgue ont découvert une propriété très importante des segments de la droite numérique. Borel a montré que si une suite infinie d'intervalles constitue un recouvrement du segment (i.e. chaque point du segment appartient à l'un au moins de ces intervalles), il existe un ensemble fini d'intervalles de cette suite qui constitue un recouvrement de ce segment. Lebesgue a établi que cette affirmation est vraie non seulement pour les suites, mais aussi pour toute famille infinie d'intervalles recouvrant le segment. Il s'est avéré très important pour les études topologiques que la proposition de Lebesgue reste valable si l'on remplace les intervalles par des ouverts quelconques, pourvu que leur réunion contienne le segment donné. On a ainsi le théorème suivant.

THÉORÈME 1 (théorème de Borel-Lebesgue). *Si un segment de la droite numérique est contenu dans la réunion d'une famille infinie* $(O_\alpha)_{\alpha \in A}$ *d'ensembles ouverts, il est encore contenu dans la réunion d'une famille finie extraite de cette famille.*

Démonstration. Pour $b = a$, le théorème est trivial. Soit $a < b$. On dira que le point x est accessible si $x \in [a, b]$ et le segment $[a, x]$ est recouvert par une famille finie d'ensembles O_α. Soit E l'ensemble de tous les points accessibles. Montrons que E vérifie les conditions du principe de l'induction continue. Comme $a \in [a, b]$, il existe un indice $\alpha \in A$ tel que $a \in O_\alpha$, d'où $a \in E$. Ainsi donc, la première condition est vérifiée. Soit maintenant $x \in E$, i.e. supposons qu'il existe une famille finie d'indices $(\alpha_i)_{i \in [1,n]}$ telle que $[a, x] \subset \bigcup_{i=1}^{n} O_{\alpha_i}$ et $x \in O_{\alpha_n}$. Si $x < b$, il existe un $x' \in]x, b]$ tel que $[x, x'] \subset O_{\alpha_n}$ puisque O_{α_n} est un voisinage du point x. Par conséquent, $[x, x''] \subset O_{\alpha_n}$ pour tout $x'' \in [x, x']$. On a alors $[a, x''] = [a, x] \cup [x, x''] \subset \bigcup_{i=1}^{n} O_{\alpha_i}$ pour tout $x'' \in [x, x']$, i.e. $[x, x'] \subset E$.

*) Le lecteur aura besoin de ce paragraphe au chapitre 8.

Enfin, soient $x > a$ et $[a, x[\subset E$ (il existe un tel x puisque $a \in E$ et la deuxième condition est vérifiée). Comme $x \in [a, b]$, $x \in O_{\alpha_0}$ pour un indice α_0. Puisque O_{α_0} est un voisinage du point x et $x > a$, il existe un x' tel que $x' \in [a, x[$ et $[x', x] \subset O_{\alpha_0}$. Comme $x' \in [a, x[$ et $[a, x[\subset E$, le point x' est accessible, i.e. il existe une famille finie d'indices $(\alpha_i)_{i \in [1,n]}$ telle que $[a, x'] \subset \bigcup_{i=1}^{n} O_{\alpha_i}$. Il s'ensuit que $[a, x] = [a, x'] \cup [x', x] \subset \bigcup_{i=0}^{n} O_{\alpha_i}$, i.e. $x \in E$. Donc, E satisfait à toutes les conditions du principe de l'induction continue, d'où $b \in E$, ce qui démontre le théorème 1.

EXERCICE 1. Montrer que les segments sont les seuls intervalles non vides dans **R** possédant la propriété affirmée par le théorème 1.

2. Ensemble compact

C'est à Alexandrov qu'appartient le mérite de généraliser le théorème de Borel-Lebesgue à tous les espaces topologiques et d'introduire la nouvelle notion de compacité (ou *bicompacité* *)) d'après sa terminologie).

DÉFINITION 1. Un ensemble C de l'espace topologique X est dit *compact* si tout recouvrement ouvert de C dans X (i.e. toute famille d'ouverts dans X constituant un recouvrement de C) possède un *sous-recouvrement fini* (i.e. une sous-famille finie constituant un recouvrement de C) ; X est appelé *espace compact* s'il est son sous-ensemble compact, i.e. si tout recouvrement ouvert de X possède un sous-recouvrement fini.

Le théorème 1 montre que *les segments sont des ensembles compacts de* **R**.

THÉORÈME 2. *Soient X un espace topologique, E son sous-espace et $C \subset E$. C est un ensemble compact de X si et seulement si C est un ensemble compact de E.*

Démonstration. Soit C un ensemble compact de E et soit $(O_\alpha)_{\alpha \in A}$ une famille d'ouverts de X constituant un recouvrement de C. Pour tout $\alpha \in A$, posons $V_\alpha = O_\alpha \cap E$. En vertu du théorème 1.4.5, V_α sont des ensembles ouverts de E. Comme $C \subset \bigcup_{\alpha \in A} O_\alpha$, on a $C \subset \left(\bigcup_{\alpha \in A} O_\alpha \right) \cap E =$

*) A l'époque, on disait que l'ensemble K d'un espace topologique est compact si toute partie infinie de K possède un point d'accumulation appartenant à K (il est à noter que le théorème de compacité (ainsi définie) du segment numérique, ou la proposition équivalente à ce théorème, s'appelait théorème de Bolzano-Weierstrass et remplaçait souvent le théorème de Borel-Lebesgue du cours des raisonnements). Cependant, la bicompacité s'est avérée une propriété beaucoup plus importante et les mathématiciens ont pris le terme « compacité » tout court (en adoptant le terme d'ensemble « dénombrablement compact » pour un ensemble compact au sens ancien). D'ailleurs les adeptes de l'école d'Alexandrov utilisent toujours l'ancienne terminologie.

$= \bigcup_{\alpha \in A} (O_\alpha \cap E) = \bigcup_{\alpha \in A} V_\alpha$. Donc, $(V_\alpha)_{\alpha \in A}$ est une famille d'ouverts de E constituant un recouvrement de C et, par suite, $C \subset \bigcup_{\alpha \in K} V_\alpha$ où K est un ensemble fini d'indices α. On a alors, à plus forte raison, $C \subset \bigcup_{\alpha \in K} O_\alpha$. On peut donc conclure que tout recouvrement ouvert de l'ensemble C dans X possède un sous-recouvrement fini, i.e. C est un ensemble compact de X. Inversement : soient C un ensemble compact de X et $(V_\alpha)_{\alpha \in A}$ une famille d'ouverts de E recouvrant C. En vertu du théorème 1.4.5, il existe pour tout $a \in A$ le plus grand ensemble parmi les ouverts de X dont la trace sur E est V_α ; notons-le O_α. Comme $(O_\alpha)_{\alpha \in A}$ est une famille d'ouverts de X recouvrant C, on a $C \subset \bigcup_{\alpha \in K} O_\alpha$ où K est un ensemble fini d'indices α. Mais alors $C \subset \bigcup_{\alpha \in K} V_\alpha$. Ainsi donc, tout recouvrement ouvert de l'ensemble C dans E possède un sous-recouvrement fini, i.e. C est un ensemble compact de E.

COROLLAIRE. *Soient X un espace topologique et $C \subset X$. C est un ensemble compact de X si et seulement si C, muni d'une topologie induite par celle de X, est un espace compact.*

En effet, prenons, dans le théorème 2 en qualité de E, le sous-espace C muni de la topologie mentionnée. D'après le théorème 2, C est un ensemble compact de X si et seulement si C est un ensemble compact de l'espace C, i.e. si C est un espace compact.

Ainsi, en particulier, *les segments de* **R** *munis d'une topologie induite sont des espaces compacts.*

Citons la définition « duale » de la compacité. Une famille d'ensembles $(E_\alpha)_{\alpha \in A}$ est dite *centrée* si chacune de ses sous-familles finies a une intersection non vide. Ainsi, chaque suite décroissante d'ensembles non vides est centrée parce que l'intersection de toute famille finie de ces ensembles est celui d'entre eux qui a le plus grand indice, et donc n'est pas vide.

THÉORÈME 3. *Un espace topologique X est compact si et seulement si chaque famille centrée de ses sous-ensembles fermés a une intersection non vide.*

Démonstration. Soit $(F_\alpha)_{\alpha \in A}$ une famille centrée d'ensembles fermés de X. Posons $O_\alpha = X \setminus E_\alpha$ pour tout $\alpha \in A$. Aucune sous-famille finie $(O_{\alpha_k})_{k \in [1,n]}$ de la famille $(O_\alpha)_{\alpha \in A}$ ne recouvre tout le X puisque $X \setminus \left(\bigcup_{k=1}^{n} O_{\alpha_k}\right) = \bigcap_{k=1}^{n} F_{\alpha_k} \neq \varnothing$. Or O_α sont des ensembles ouverts. C'est pourquoi, si X est compact, toute la famille $(O_\alpha)_{\alpha \in A}$, elle non plus, ne

recouvre pas X, ce qui implique $\bigcap_{\alpha \in A} \dot{F}_\alpha = X \setminus \left(\bigcup_{\alpha \in A} O_\alpha \right) \neq \varnothing$. Ainsi donc, si X est compact, toute famille centrée de ses sous-ensembles fermés a une intersection non vide. Inversement : supposons que X possède cette dernière propriété et que $(O_\alpha)_{\alpha \in A}$ soit une famille d'ouverts recouvrant X. Posons $F_\alpha = X \setminus O_\alpha$ pour tout $\alpha \in A$. Comme $\bigcap_{\alpha \in A} F_\alpha = X \setminus \bigcup_{\alpha \in A} O_\alpha = \varnothing$ et $(F_\alpha)_{\alpha \in A}$ est une famille d'ensembles fermés, celle-ci n'est pas centrée, i.e. elle possède une sous-famille finie $(F_{\alpha_k})_{k \in [1,n]}$ qui a une intersection vide. On a alors $\bigcup_{k=1}^{n} O_{\alpha_k} = X \setminus \bigcap_{k=1}^{n} F_{\alpha_k} = X$, de sorte que tout recouvrement ouvert de l'espace X possède un sous-recouvrement fini, i.e. X est compact.

COROLLAIRE. *Toute suite décroissante d'ensembles fermés non vides de l'espace compact a une intersection non vide.*

REMARQUE. Il existe dans **R** des suites décroissantes d'ensembles fermés non vides ayant une intersection vide, par exemple, $([n, +\infty[)_{n \in \mathbf{N}}$. D'où l'on déduit que **R** n'est pas compact.

THÉORÈME 4. *La réunion C d'une famille finie $(C_i)_{i \in I}$ de sous-ensembles compacts de l'espace topologique X est un ensemble compact.*

Démonstration. Soit $(O_\alpha)_{\alpha \in A}$ une famille d'ouverts de X recouvrant C et, par conséquent, tout C_i. Etant donné que C_i sont des ensembles compacts de X, il existe des ensembles finis $K_i \subset A$ $(i \in I)$ tels que $C_i \subset \bigcup_{\alpha \in K_i} O_\alpha$. Alors $C \subset \bigcup_{\alpha \in K} O_\alpha$ où $K = \bigcup_{i \in I} K_i$. Comme I et tous les K_i sont finis, K est également fini. Ainsi donc, tout recouvrement ouvert de l'ensemble C dans X possède un recouvrement fini, i.e. C est un ensemble compact.

COROLLAIRE. *Tout sous-ensemble fini de l'espace topologique est compact.*

En effet, il représente la réunion de la famille finie des singletons et ces derniers sont évidemment compacts.

THÉORÈME 5. *Tout sous-ensemble fermé F de l'espace compact X est compact.*

Démonstration. Comme F est fermé, $X \setminus F$ est ouvert. En le joignant à n'importe quelle famille S d'ouverts de l'espace X recouvrant F, on obtient un recouvrement ouvert de tout l'espace X. Puisque X est compact, ce recouvrement possède un sous-recouvrement fini. En éliminant $X \setminus F$ de ce dernier (s'il y entre), on obtient une sous-famille finie de S recouvrant F. Ainsi donc, chaque recouvrement ouvert S de F dans X possède un sous-recouvrement fini, i.e. F est un ensemble compact de X.

THÉORÈME 6. *Soient X et Y des espaces topologiques, f une application de X dans Y et C un ensemble compact de X contenu dans D_f. Si la restriction $f|_C$ de f à C (C étant muni de la topologie induite par celle de X) est continue, $f(C)$ est un ensemble compact de Y.*

Démonstration. Soit $(V_\alpha)_{\alpha \in A}$ un recouvrement ouvert de l'ensemble $f(C)$ dans Y. Comme $f|_C$ est continue, le théorème 2.2.3 dit que $U_\alpha = f|_C^{-1}(V_\alpha)$ sont des ouverts de C et, de cette façon, $(U_\alpha)_{\alpha \in A}$ est un recouvrement ouvert de l'espace C. Puisque C est un espace compact en vertu du théorème 2, on a $C = \bigcup_{\alpha \in K} U_\alpha$, où K est une partie finie de l'ensemble A. Il s'ensuit que $f(C) \subset \bigcup_{\alpha \in K} f(U_\alpha) \subset \bigcup_{\alpha \in K} V_\alpha$, de sorte que tout recouvrement ouvert de $f(C)$ dans Y possède un sous-recouvrement fini, i.e. $f(C)$ est un ensemble compact de Y.

COROLLAIRE. *L'ensemble des points de l'arc de Jordan est compact.*

En effet, il a été noté que les segments dans $\mathbf{R}$ sont compacts et que l'ensemble des points de l'arc de Jordan est l'image d'un segment par une application continue de $\mathbf{R}$ dans $\mathbf{R}^n$.

3. Fonction semi-continue

DÉFINITION 2. On dit qu'une fonction f de l'espace topologique X dans $\bar{\mathbf{R}}$ est *semi-continue inférieurement* (resp. *semi-continue supérieurement*) *en un point* x, si $x \in D_f$ et pour tout $c < f(x)$ (resp. pour tout $d > f(x)$) il existe un voisinage U du point x tel que $c < f(U)$ (resp. $d > f(U)$). La fonction f est dite *semi-continue inférieurement* (resp. *semi-continue supérieurement*) si $D_f = X$ et f est semi-continue inférieurement (resp. supérieurement) en tout point $x \in X$.

EXEMPLE 1. Soit χ_E une fonction caractéristique du sous-ensemble E de l'espace topologique X (i.e. une fonction sur X qui est égale à l'unité sur E et à zéro sur $X \setminus E$). *Si E est ouvert* (*resp. fermé*), *χ_E est semi-continue inférieurement* (*resp. supérieurement*). En effet, soit E un ouvert. Si $x \in E$ (de sorte que $\chi_E(x) = 1$ et E est un voisinage du point x), on a $\chi_E(E)$ $(= \{1\}) > c$ pour tout $c < \chi_E(x)$. Si $x \in X \setminus E$ (de sorte que $\chi_E(x) = 0$), on obtient que même $\chi_E(X)$ $(\geqslant 0) > C$ pour tout $C < \chi_E(x)$. Soit maintenant E un ensemble fermé. Si $x \in E$ (de sorte que $\chi_E(x) = 1$), on a $\chi_E(X)$ $(\leqslant 1) < d$ pour tout $d > \chi_E(x)$. Si $x \in X \setminus E$ (de sorte que $X \setminus E$ est un voisinage du point x et $\chi_E(x) = 0$), il vient $\chi_E(X \setminus E)$ $(= \{0\}) < d$ pour tout $d > \chi_E(x)$.

EXERCICE 2. Démontrer que la réciproque est aussi vraie : si la fonction caractéristique de l'ensemble E est semi-continue inférieurement (resp. supérieurement), E est ouvert (resp. fermé).

THÉORÈME 7. *Une fonction f de l'espace topologique X dans* $\mathbf{R}$ *est continue en un point $x \in X$ si et seulement si f (en tant que fonction de X dans $\bar{\mathbf{R}}$) est semi-continue en x à la fois inférieurement et supérieurement.*

Démonstration. Soit f une fonction semi-continue en x inférieurement et supérieurement (d'où, en particulier, il découle que $x \in D_f$). Il existe alors pour tout nombre $\varepsilon > 0$ des voisinages U_1 et U_2 du point x tels que $f(x') > f(x) - \varepsilon$ pour tout $x' \in U_1 \cap D_f$, et $f(x') < f(x) + \varepsilon$ pour tout $x' \in U_2 \cap D_f$, d'où $|f(x') - f(x)| < \varepsilon$ pour tout $x' \in U_1 \cap U_2 \cap D_f$. Puisque $U_1 \cap U_2$ est un voisinage du point x, cela signifie que f est continue en x. Inversement : soit f une fonction continue en x et soient c et d des points arbitraires de $\bar{\mathbf{R}}$ tels que $c < f(x) < d$. Comme $]c, d[$ est un voisinage du point $f(x)$, il existe un voisinage U du point x tel que $c < f(U) < d$. Mais cela signifie que f est semi-continue en x inférieurement et supérieurement.

Il découle de la définition 2 qu'*une fonction f de l'espace topologique X dans $\bar{\mathbf{R}}$ est semi-continue inférieurement en un point x si et seulement si $-f$ est semi-continue supérieurement en ce point.* En effet, soit $c < f(x)$, i.e. $-f(x) < -c$. Si f est semi-continue supérieurement au point x, il existe un voisinage U de ce point tel que $(-f)(U) < -c$, i.e. $c < f(U)$. Par conséquent, f est semi-continue inférieurement en x. On démontre de façon analogue la proposition réciproque.

LEMME. *Une fonction f à valeurs dans $\bar{\mathbf{R}}$, définie sur l'espace topologique X est semi-continue inférieurement (resp. supérieurement) si et seulement si $f^{-1}(]c, +\infty])$ (resp. $f^{-1}([-\infty, d[)$ est un ouvert de X pour tout $c \in \bar{\mathbf{R}}$ (resp. $d \in \bar{\mathbf{R}}$).*

Démonstration. Soit f une fonction semi-continue inférieurement. Si $x \in f^{-1}(]c, +\infty])$, i.e. $f(x) > c$, il existe un voisinage U du point x tel que $f(U) > c$, d'où $U \subset f^{-1}(]c, +\infty])$, de sorte que $f^{-1}(]c, +\infty])$ est un voisinage de x. Ainsi donc, $f^{-1}(]c, +\infty])$ est un voisinage de chacun de ses points, i.e. un ensemble ouvert. Inversement : supposons que $f^{-1}(]c, +\infty])$ soit un ouvert pout tout $c \in \mathbf{R}$ et que $c < f(x)$, de sorte que $x \in f^{-1}(]c, +\infty])$. Alors $U = f^{-1}(]c, +\infty])$ est un voisinage du point x et $f(U) \subset]c, +\infty]$, i.e. $f(U) > c$. De cette façon, f est semi-continue inférieurement en chaque point $x \in X$. Enfin, comme

$$f^{-1}([-\infty, d[) = (-f)^{-1}(]-d, +\infty]),$$

l'ouverture des ensembles $f^{-1}([-\infty, d[)$ est équivalente, d'après ce qu'on a déjà démontré, à la semi-continuité inférieure de la fonction $-f$, i.e. à la semi-continuité supérieure de la fonction f.

THÉORÈME 8. *Une fonction f à valeurs dans $\bar{\mathbf{R}}$, définie et semi-continue inférieurement (resp. supérieurement) sur un espace compact non vide X présente sur X sa valeur minimale (resp. maximale).*

Démonstration. Soient f une fonction semi-continue inférieurement et $m = \inf f (= \inf f(X))$. Montrons que m est une valeur de f. Supposons que cette affirmation ne soit pas vraie. Comme $f(X) \neq \varnothing$, on a $m < +\infty$. Posons

$$m_n = \begin{cases} m + \dfrac{1}{n} & \text{si} \quad m \text{ est fini ;} \\ -n & \text{si} \quad m = -\infty. \end{cases}$$

Comme $m_n \to m$, la suite des ensembles $(f^{-1}(]m_n, +\infty]))$ recouvre X. En effet, soit x un point arbitraire de X. Comme $m \neq f(x)$ par hypothèse, $m < f(x)$ et pour cette raison il existe un numéro n tel que $m_n < f(x)$, i.e. $x \in f^{-1}(]m_n, +\infty])$. Mais en vertu du lemme, $f^{-1}(]m_n, +\infty])$ sont des ensembles ouverts de X. Par conséquent, puisque X est compact, il existe une famille finie de ces ensembles qui recouvre X et, comme (m_n) est une suite décroissante et $(f^{-1}(]m_n, +\infty])$ est alors une suite croissante, il existe un numéro n tel que $X = f^{-1}(]m_n, +\infty])$, i.e. $f(x) > m_n$ pour tout $x \in X$. On a alors $m = \inf\limits_{x \in X} f(x) \geqslant m_n$, tandis que $m_n > m$. La contradiction obtenue montre que $m \in f(X)$ et, donc, m est la valeur minimale de f sur X. Il en découle déjà que la deuxième affirmation du théorème est aussi vraie. En effet, si f est semi-continue supérieurement et $M = \sup\limits_X f$, la fonction $-f$ est semi-continue inférieurement et $-M = \inf(-f)$; il s'ensuit d'après ce qu'on a déjà démontré que $-M \in (-f)(-X) = -f(X)$, i.e. $M \in f(X)$. D'où M est la valeur maximale de f sur X.

Les théorèmes 7 et 8 entraînent immédiatement le théorème suivant.

THÉORÈME 9. *Une fonction réelle continue sur un espace compact possède la plus petite et la plus grande valeur.*

En effet, elle possède la plus petite valeur en tant que fonction semi-continue inférieurement, et la plus grande valeur en tant que fonction semi-continue supérieurement.

COROLLAIRE. *Une fonction réelle continue sur un espace compact est bornée.*

4. Sous-ensembles compacts d'un espace séparé

Rappelons qu'un espace topologique est dit *séparé* si ses deux points distincts quelconques possèdent des voisinages disjoints. En particulier, nous avons déjà montré au n° 2.1.4. qu'un espace métrique est séparé s'il est muni de la topologie engendrée par sa métrique. Notons que le point a d'un espace séparé est séparé de tout autre point b de cet espace par un voisinage fermé. En effet, a et b possèdent des voisinages disjoints U et V ; V contient un voisinage ouvert O du point b ; il s'ensuit que $[U]$ a une intersection vide avec O et, par conséquent, ne contient pas b.

THÉORÈME 10. *Le sous-ensemble compact d'un espace topologique séparé est fermé.*

Démonstration. Supposons que cela n'est pas vrai, i.e. il existe un espace séparé X possédant un sous-ensemble compact C non fermé. Il s'ensuit que C possède un point adhérent c qui ne lui appartient pas. On en conclut que les complémentaires $X \setminus V$ dans X des voisinages fermés V de c constituent un recouvrement de l'ensemble C. En effet, si $x \in C$, il existe un voisinage fermé V de c qui ne contient pas x puisque $x \neq c$ et X est séparé, d'où l'on déduit que $x \in X \setminus V$. Comme les ensembles $X \setminus V$ sont ouverts et C est compact, il existe une famille finie $(V_i)_{i \in [1,n]}$ de voisinages fermés du point c telle que $C \subset \bigcup_{i=1}^{n} (X \setminus V_i) = X \setminus \bigcap_{i=1}^{n} V_i$, i.e. le voisinage $\bigcap_{i=1}^{n} V_i$ du point c a une intersection vide avec l'ensemble C, ce qui contredit la relation $c \in [C]$.

REMARQUE. Les sous-ensembles compacts d'un espace non séparé peuvent être non fermés. Ainsi, les singletons de tout espace sont des ensembles compacts, mais si, par exemple, X est un espace non vide contenant plus d'un élément, dont les seuls ensembles fermés sont $\varnothing$ et X, les singletons de X ne sont pas fermés.

THÉORÈME 11. *L'application g réciproque d'une application injective continue f d'un espace compact C dans un espace topologique séparé X est continue si $D_g = f(C)$ est muni de la topologie induite par celle de X.*

Démonstration. Soit F un fermé arbitraire de C. En vertu du théorème 5, F est un ensemble compact de C et, par conséquent, en vertu du théorème 6, $f(F)$ est un ensemble compact de X. Comme X est séparé, $f(F)$ est fermé dans X suivant le théorème 10 et, par conséquent, selon le théorème 1.4.5 il est aussi fermé dans l'espace $f(C)$ muni de la topologie induite par celle de X. Or $g^{-1}(F) = f(F)$. Par suite, l'image réciproque par g de tout ensemble fermé de C est fermée dans $f(C)$, donc, d'après le théorème 2.2.3′, l'application g est continue.

Notons encore deux propriétés des sous-ensembles compacts de l'espace séparé.

1° *L'intersection de toute famille non vide de sous-ensembles compacts de l'espace séparé est compacte.* En effet, en tant qu'une intersection d'ensembles fermés (voir théorème 10), elle est fermée (théorème 1.4.1′) et, comme sous-ensemble fermé d'ensembles compacts de la famille, elle est compacte (voir théorème 5).

REMARQUE. L'intersection de deux sous-ensembles compacts de l'espace non séparé peut déjà ne pas être compacte.

2° *Une famille centrée de sous-ensembles compacts de l'espace séparé (en particulier, la suite décroissante de ses sous-ensembles compacts non*

vides) a une intersection non vide. En effet, les intersections de tous ces sous-ensembles avec l'un quelconque d'entre eux sont compactes d'après la propriété 1° et, par conséquent, suivant le théorème 10, sont fermées. En outre, ces intersections forment, elles aussi, une famille centrée. Donc, cette famille possède, selon le théorème 3, une intersection non vide, et cette dernière est, évidemment, l'intersection de la famille centrée initiale.

5. Ensembles compacts d'un espace métrique. Distance à un ensemble

DÉFINITION 3. Un ensemble E d'un espace métrique X est dit *borné* s'il existe une boule qui le contient, c'est-à-dire si l'ensemble des distances de ses points à un point donné de X est borné.

Si E est borné, l'ensemble des distances de ses points à n'importe quel point donné de X est borné. En effet, si $d(x, x_0) < C$ pour tout $x \in E$, et x_1 est un point arbitraire de l'espace X, on a $d(x, x_1) \leqslant d(x, x_0) + d(x_0, x_1) < C + d(x_0, x_1)$ pour tout $x \in E$. Si X est un espace normé (en particulier, $\mathbf{R}^n$), il est plus commode de prendre le point 0 pour x_0 ; comme $d(x, 0) = \|x\|$, *l'ensemble E de X est borné si et seulement s'il est « borné pour la norme », i.e. l'ensemble numérique $\{\|x\| \mid x \in E\}$ est borné.*

Nous allons appeller *compacts* les sous-ensembles compacts des espaces métriques et, en particulier, les espaces métriques compacts.

THÉORÈME 12. *Les compacts sont bornés.*

Démonstration. Soit E un compact, i.e. un sous-ensemble compact d'un espace métrique X. La suite des boules ouvertes $(U(x_0\,;\,n))_{n \in \mathbf{N}}$ au centre en un point $x_0 \in X$ recouvre X et, par conséquent, E aussi. Comme E est un compact et cette suite est strictement croissante, $E \subset U(x_0\,;\,n)$ pour un $n \in \mathbf{N}$, ce qui signifie justement que E est borné.

DÉFINITION 4. Soient x et A un point et un ensemble de l'espace métrique X. On appelle *distance de x à A* la borne inférieure $\varrho(x, A)$ des distances de x aux points de A :

$$\varrho(x, A) = \inf_{a \in A} d(x, a).$$

On dit qu'elle est *atteinte* s'il existe un point $a_0 \in A$ tel que $\varrho(x, A) = d(x, a_0)$ et on dit alors que $d(x, a_0)$ est *la plus courte distance de x à A.*

Ainsi donc, on a toujours $\varrho(x, A) \geqslant 0$. Si $A \neq \varnothing$, $\varrho(x, A) < +\infty$ en tant que borne inférieure de l'ensemble non vide $\{\varrho(x, a) \mid a \in A\}$. Enfin, $\varrho(x, \varnothing) = \inf \varnothing = +\infty$.

Notons quelques propriétés des distances pour un ensemble non vide.

1° $|\varrho(x, A) - \varrho(y, A)| \leqslant d(x, y)$. C'est pourquoi, $\varrho(x, A)$ est une *fonction continue de x.* En effet, $d(x, a) \leqslant d(x, y) + d(y, a)$ pour tout $a \in A$. En passant à la borne inférieure par rapport à a dans le premier membre (et en fixant a dans le second), on obtient $\varrho(x, A) \leqslant d(x, y) + d(y, a)$, i.e. $\varrho(x, A) - d(x, y) \leqslant d(y, a)$. Comme le premier membre ne dépend pas

de a, le passage à la borne inférieure par rapport à a dans le second membre donne $\varrho(x, A) - d(x, y) \leqslant \varrho(y, A)$, i.e.

$$\varrho(x, A) - \varrho(y, A) \leqslant d(x, y).$$

Puisque x et y sont équivalents, on a aussi

$$\varrho(y, A) - \varrho(x, A) \leqslant d(y, x) = d(x, y).$$

Par conséquent, $|\varrho(x, A) - \varrho(y, A)| = \max(\varrho(x, A) - \varrho(y, A), \varrho(y, A) - - \varrho(x, A)) \leqslant d(x, y)$.

REMARQUE. En particulier, $d(x, x_0)$ $(= \varrho(x, \{x_0\}))$ est une *fonction continue de* x. D'ailleurs cela découle directement de l'inégalité $|d(x, x_0) - - d(y, x_0)| \leqslant d(x, y)$.

2° *La distance d'un point x à un ensemble A est atteinte si A est un compact non vide.* En effet, $d(x, a)$ en tant que fonction continue de a sur le compact A possède sur lui, d'après le théorème 9, la plus petite valeur.

Notons qu'il existe des espaces normés où la distance d'un point à un ensemble non vide fermé (et même borné) peut ne pas être atteinte.

EXERCICE 3. Montrer que la distance d'un point à un ensemble est atteinte pour tout ensemble fermé non vide de $\mathbf{R}^n$.

3° *Les ensembles*

$$B(A\,;\,\delta) = \{x \in X |\ \varrho(x, A) \leqslant \delta\} \qquad (\delta \in [0, +\infty[)$$

sont fermés. Ceci étant, $B(A\,;\,0) = [A]$, *de sorte que si A est fermé, on a* $\varrho(x, A) = 0 \Rightarrow x \in A$. *Si A est borné, $B(A\,;\,\delta)$ est aussi borné.* En effet, soit $x \in [B(A\,;\,\delta)]$. Il existe alors pour tout $\varepsilon > 0$ un $x' \in B(A\,;\,\delta)$ tel que $d(x, x') < \varepsilon$. Comme $x' \in B(A\,;\,\delta)$, i.e. $\inf\limits_{a \in A} d(x', a) \leqslant \delta$, il existe pour ce ε un $a \in A$ tel que $d(x', a) < \delta + \varepsilon$. On a alors $\varrho(x, A) \leqslant d(x, a) \leqslant \leqslant d(x, x') + d(x', a) < \delta + 2\varepsilon$ et, comme $\varrho(x, A)$ ne dépend pas de ε, il vient $\varrho(x, A) \leqslant \delta$, i.e. $x \in B(A\,;\,\delta)$. Ainsi donc, $[B(A\,;\,\delta)] \subset B(A\,;\,\delta)$, i.e. $B(A\,;\,\delta)$ est fermé. On a par ailleurs $\inf\limits_{a \in A} d(x, a) = 0$, i.e. $x \in B(A\,;\,0)$, si et seulement si pour tout $\varepsilon > 0$ il existe un $a \in A$ tel que $d(x, a) < \varepsilon$, i.e. si $U(x, \varepsilon) \cap A \neq \varnothing$ pour tout $\varepsilon > 0$; mais cela est équivalent à $x \in [A]$. Soit maintenant A un ensemble borné, i.e. il existe en vertu de la définition 3 un point $x_0 \in X$ et un nombre $r > 0$ tels que $d(a, x_0) < r$ pour tout $a \in A$. Si $x \in B(A\,;\,\delta)$, i.e. $\varrho(x, A) \leqslant \delta$, on a $d(x, a) < \delta + 1$ pour un point $a \in A$ et, pour cette raison, $d(x, x_0) \leqslant d(x, a) + d(a, x_0) < \delta + 1 + r$; il s'ensuit que $B(A\,;\,\delta) \subset U(x_0\,;\,\delta + 1 + r)$, de sorte que $B(A\,;\,\delta)$ est borné.

DÉFINITION 5. On appelle *distance entre les ensembles A et B* de l'espace métrique la borne inférieure $\varrho(A, B)$ des distances des points de l'ensemble A aux points de l'ensemble B :

$$\varrho(A, B) = \inf\{d(a, b) |\ a \in A,\ b \in B\}.$$

On dit qu'elle est *atteinte* s'il existe des points $a_0 \in A$ et $b_0 \in B$ tels que $\varrho(A, B) = d(a_0, b_0)$ et on dit alors que $d(a_0, b_0)$ est *la plus courte distance entre les ensembles A et B.*

Ainsi donc, $\varrho(A, B) \geqslant 0$. Ceci étant, $\varrho(A, B) < +\infty$ si et seulement si A et B ne sont pas vides.

4° *La distance entre deux ensembles est atteinte s'ils sont des compacts non vides.* En effet, soient A et B des compacts non vides de l'espace métrique X. D'après la propriété 1°, $\varrho(A, B)$ est continue sur A, de sorte qu'elle prend en un point $a_0 \in A$ la plus petite valeur (voir théorème 9). Il s'ensuit que $\varrho(a_0, B) \leqslant \varrho(a, B)$ pour tout $a \in A$ et, par conséquent, $\varrho(a_0, B) \leqslant \leqslant d(a, b)$ pour tout $a \in A$ et tout $b \in B$. D'autre part, en vertu de la propriété 2° il existe un point $b_0 \in B$ pour lequel $\varrho(a_0, B) = d(a_0, b_0)$. Ainsi donc, $d(a_0, b_0) \leqslant d(a, b)$ pour tous $a \in A$ et $b \in B$, i.e. $d(a_0, b_0)$ est la plus courte distance entre les ensembles A et B.

5° *Si la distance entre un compact A et un ensemble fermé B est égale à zéro, ces deux ensembles ont une intersection non vide.* Ainsi donc, *la distance entre deux ensembles disjoints dont l'un est un compact et l'autre un fermé (en particulier, entre deux compacts disjoints) est strictement positive.* En effet, comme $\varrho(A, B) < +\infty$, A et B ne sont pas vides. Vu que $\varrho(a, B)$ est une fonction continue de a (propriété 1°), elle atteint en un point a_0 du compact A la plus petite valeur (voir théorème 9). Ainsi donc, $\varrho(a_0, B) \leqslant \varrho(a, B)$ pour tout $a \in A$ et, par conséquent, $\varrho(a_0, B) \leqslant d(a, b)$ pour tous $a \in A$ et $b \in B$. On obtient alors $0 \leqslant \varrho(a_0, B) \leqslant \varrho(A, B) = 0$, i.e. $\varrho(a_0, B) = 0$. Comme B est fermé, il en découle en vertu de la propriété 3° que $a_0 \in B$, de sorte que $A \cap B \neq \varnothing$.

REMARQUE. La distance entre deux fermés disjoints peut être égale à zéro. C'est le cas, par exemple, d'une branche de l'hyperbole et de son asymptote.

EXERCICE 4. Construire un exemple de deux fermés disjoints de **R**, la distance entre lesquels est égale à zéro

6. Compacts de $\mathbf{R}^n$

Le théorème suivant est une généralisation directe du théorème 1 de Borel-Lebesgue.

THÉORÈME 1′. *Les parallélépipèdes*

$$\Pi = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid x_k \in [a_k, b_k] \ (k = 1, \ldots, n)\} \qquad (1)$$

(i.e. les produits cartésiens $[a_1, b_1] \times \ldots \times [a_n, b_n]$*) où* $[a_1, b_1], \ldots, [a_n, b_n]$ *sont des segments arbitraires dans* **R**, *sont des compacts de* $\mathbf{R}^n$.

Démonstration. On procède par récurrence sur n. Pour $n = 1$ (où Π est un segment $[a_1, b_1]$ dans **R**) c'est le cas du théorème 1 de Borel-Lebesgue. Supposons que le théorème soit vrai pour les parallélépipèdes dans $\mathbf{R}^{n-1}$ où $n > 1$. Soit Π un parallélépipède (1) dans $\mathbf{R}^n$ et soit $(O_\alpha)_{\alpha \in A}$ son recou-

vrement ouvert. Il s'ensuit que Π peut être présenté sous la forme

$$\Pi = \Pi' \times [a_n, b_n]$$

où Π' est un parallélépipède $[a_1, b_1] \times \ldots \times [a_{n-1}, b_{n-1}]$ dans $\mathbf{R}^{n-1}$.

1) Pour tout $t \in [a_n, b_n]$, notons $\Pi'(t)$ la section

$$\{(x_1, \ldots, x_n) \in \Pi \mid x_n = t\} = \Pi' \times \{t\}$$

du parallélépipède (1) par l'hyperplan $x_n = t$. D'après l'hypothèse de récurrence, Π' est un compact dans $\mathbf{R}^{n-1}$. Mais $\Pi'(t)$ est son image par une application $\mathbf{R}^{n-1} \to \mathbf{R}^n$ telle que

$$(x_1, \ldots, x_{n-1}) \mapsto (x_1, \ldots, x_{n-1}, t).$$

Comme elle est continue, $\Pi'(t)$ est aussi un compact d'après le théorème 6.

2) Comme $(O_\alpha)_{\alpha \in A}$ est un recouvrement ouvert du compact $\Pi'(t)$, il existe une famille finie d'indices $(\alpha_i)_{i \in [1,k]}$ telle que $\Pi'(t) \subset O$, où $O = \bigcup_{i=1}^{k} O_{\alpha_i}$. Soit $F = \mathbf{R}^n \setminus O$. Comme F est un ensemble fini ayant une intersection vide avec le compact $\Pi'(t)$, la propriété 5° démontrée au numéro précédent de ce paragraphe dit que $\varrho(F, \Pi'(t)) > 0$. On obtient alors

$$\Pi'^{\,t+h}_{t-h} = \Pi'_{t_i} \times [t - h, t + h] \subset O$$

pour tout $h \in]0, \varrho(F, \Pi'(t))[$. En effet, si $x \in \Pi'^{\,t+n}_{t-h}$, on a $\varrho(x, \Pi'(t)) < h < \varrho(F, \Pi'(t))$, de sorte que $x \notin F$ et par conséquent, $x \in O$. Donc il existe pour tout $t \in [a_n, b_n]$ un nombre $h > 0$ tel que $\Pi'^{\,t+h}_{t-h}$ est recouvert par une famille finie d'ouverts O_α.

3) Soit E_t l'ensemble de tous les $h > 0$ pour lesquels $\Pi'^{\,t+h}_{t-h}$ est recouvert par une famille finie d'ouverts. D'après le n° 2, $E_t \neq \varnothing$ et donc $\sup E_t > 0$. S'il existe un t tel que $\sup E_t > b_n - a_n$, on trouve dans E_t un nombre $h > b_n - a_n$, de sorte que $\Pi'^{\,t+h}_{t-h}$, et avec lui $\Pi'^{\,t+(b_n-a_n)}_{t-(b_n-a_n)}$, est recouvert par une famille finie d'ouverts. Mais comme $t \in [a_n, b_n]$, on a $t - (b_n - a_n) \leqslant a_n$ et $t + (b_n - a_n) \geqslant b_n$, de sorte que $[a_n, b_n] \subset [t - (b_n - a_n), t + (b_n - a_n)]$. Par conséquent, il existe une famille finie d'ouverts qui recouvre $\Pi' \times [a_n, b_n]$, i.e. Π, c.q.f.d.

4) Soit, enfin, $\sup E_t \leqslant b_n - a_n$ pour tout $t \in [a_n, b_n]$. Les intervalles $]t - h_t, t + h_t[$, où $h_t = \sup E_t$, forment un recouvrement ouvert du segment $[a_n, b_n]$ de sorte qu'il existe d'après le théorème de Borel-Lebesgue des points $t_1, \ldots, t_l$, tels que

$$[a_n, b_n] \subset \bigcup_{i=1}^{l}]t_i - h_{t_i}, t_i + h_{t_i}[\subset \bigcup_{i=1}^{l} [t_i - h_{t_i}, t_i + h_{t_i}],$$

d'où $\Pi \subset \bigcup_{i=1}^{l} \Pi_{t_i - h_{t_i}}^{t_i + h_{t_i}}$. Puisque $h_{t_i} < \sup E_{t_i}$, il existe dans E_{t_i} un nombre $h_i > h_{t_i}$, de sorte que $\Pi_{t_i - h_i}^{t_i + h_i}$, et avec lui $\Pi_{t_i - h_{t_i}}^{t_i + h_{t_i}}$, est recouvert par une famille finie d'ouverts. Ainsi le parallélépipède Π est recouvert par une famille finie $(\Pi_{t_i - h_{t_i}}^{t_i + h_{t_i}})_{i \in [1, l]}$ d'ensembles ayant chacun un recouvrement ouvert fini. Par conséquent, il existe une famille finie d'ouverts qui recouvre Π.

Donc, il est démontré dans 3) et 4) que le recouvrement ouvert $(O_\alpha)_{\alpha \in A}$ du parallélépipède Π possède un sous-recouvrement fini. Donc, Π est un compact, et le théorème est ainsi démontré.

THÉORÈME 13. *Un ensemble F de $\mathbf{R}^n$ est un compact si et seulement s'il est borné et fermé.*

Démonstration. Le compact F de $\mathbf{R}^n$ est borné en vertu du théorème 12 et fermé en vertu du théorème 10. Inversement : soit F un ensemble borné et fermé dans $\mathbf{R}^n$. Etant borné, il est contenu dans une boule $B(0\,;\,r)$ et, par conséquent, dans un parallélépipède $\Pi = \prod_{k=1}^{n} \Delta_k$ où $\Delta_k = [-r, r]$ $(k = 1, \ldots, n)$. Etant fermé dans $\mathbf{R}^n$, il est fermé, en vertu du théorème 1.4.5, dans le sous-espace Π de l'espace $\mathbf{R}^n$. Mais, d'après le théorème 1′, Π est un compact de $\mathbf{R}^n$ et, donc, un sous-espace compact d'après le corollaire du théorème 2. Par conséquent, F est un ensemble compact dans Π d'après le théorème 5 et, donc, dans $\mathbf{R}^n$, d'après le théorème 2.

EXEMPLE 2. *Si A est un ensemble borné de $\mathbf{R}^n$, les ensembles*

$$B(A\,;\,\delta) = \{x \in \mathbf{R}^n \mid \varrho(x, A) \leqslant \delta\} \quad (\delta \in [0, +\infty[)$$

sont des compacts. En effet, en vertu de la propriété 3° des distances démontrée au n° précédent de ce paragraphe, les ensembles $B(A\,;\,\delta)$ sont bornés et fermés. Par conséquent, d'après le théorème 13, ils sont compacts. En particulier, *si A est un compact de $\mathbf{R}^n$, tous les $B(A\,;\,\delta)$ sont des compacts* puisque le compact A est borné.

Comme $B(x_0\,;\,r) = B(\{x_0\}\,;\,r)$, il résulte de ce que nous venons de dire que *les boules fermées $B(x_0\,;\,r)$ de $\mathbf{R}^n$ sont des compacts.*

EXEMPLE 3. *Les sphères $S(x_0\,;\,r) = \{x \in \mathbf{R}^n \mid d(x_0, x) = r\}$ de $\mathbf{R}^n$ sont des compacts.* En effet, d'après leur définition, elles sont bornées. Et comme

$$S(x_0\,;\,r) = B(x_0\,;\,r) \setminus U(r_0\,;\,r)$$

et $U(x_0\,;\,r)$ sont des ensembles ouverts (voir théorème 1.3.1), les sphères $S(x_0\,;\,r)$ avec les boules $B(x_0\,;\,r)$ sont fermées.

7. Continuité uniforme

Dans le n° 3 il a été établi (théorème 9) que toute fonction réelle conti-

nue sur un espace compact présente la plus petite et la plus grande valeur. Dans le cas d'espaces compacts métriques, les fonctions réelles continues (et, d'une façon plus générale, les applications continues dans les espaces métriques) possèdent encore une propriété très importante : continuité uniforme. (Vu que nous en aurons besoin pour la première fois dans le chapitre 11, le lecteur peut remettre ce sujet au plus tard.)

DÉFINITION 6. Soient X et Y des espaces métriques munis de métriques d_X et d_Y respectivement. Une application f de X dans Y est dite *continue sur un ensemble E* si $E \subset D_f$ et la restriction de f à E est continue, i.e.

$$(2) \quad (\forall x \in E)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x' \in E)\,(d_X(x, x') < \delta \Rightarrow \Rightarrow d_Y(f(x), f(x')) < \varepsilon).$$

Il est évident que la continuité de f en tout point $x \in E$ entraîne celle de f sur E. Mais la réciproque n'est pas vraie. Ainsi, la fonction de Dirichlet $D(x)$ qui est égale à 1 si x est rationnel et à 0 si x est irrationnel et dont l'ensemble de définition est tout le $\mathbf{R}$, est discontinue en tout point $x \in \mathbf{R}$. Cependant, elle est continue sur tout ensemble E de nombres rationnels (ainsi que sur tout ensemble de nombres irrationnels) car sur ces ensembles elle est constante.

DÉFINITION 7. Soient X et Y des espaces métriques munis de métriques d_X et d_Y respectivement. Une application f de X dans Y est dite *uniformément continue sur un ensemble E* si $E \subset D_f$ et

$$(3) \quad (\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in E)(\forall x' \in E)\,(d_X(x, x') < \delta \Rightarrow \Rightarrow d_Y(f(x), f(x')) < \varepsilon),$$

i.e. quels que soient les points x et x' de E, la distance entre les valeurs de f en ces points est strictement inférieure à ε dès que la distance entre x et x' est strictement inférieure à δ.

Il est évident qu'une application uniformément continue sur l'ensemble E est uniformément continue sur chacun des sous-ensembles de E.

La continuité uniforme entraîne la continuité simple. En effet, en fixant $x \in E$, on obtient de la condition (3) que

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x' \in E)(d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon),$$

et puisque x est arbitraire dans E, c'est la même chose que (2). Qu'est-ce que la condition (3) ajoute à la condition (2) ? Elle ajoute ce que pour tout ε donné il existe un δ valable en même temps pour tout $x \in E$. C'est ce qu'on appelle « uniformité ». Dans la condition (2), δ dépend non seulement de ε, mais de x aussi. Il est évident que la donnée de ε et x ne définit

pas δ univoquement : si un $\delta > 0$ satisfait à la condition (2) pour ε et x donnés, tout $\delta' > 0$ inférieur à δ satisfait à cette condition pour les mêmes ε et x. Cependant, on ne peut pas le dire pour tout δ aussi grand que l'on veut. Il s'avère que parmi les valeurs « admissibles » de δ, il existe la plus grande (probablement, égale à $+\infty$) qui est définie univoquement par ε et x donnés. Plus précisément : pour tout $x \in E$ et tout $\varepsilon > 0$ posons

$$C_{x,\varepsilon} = \{\delta \in]0, +\infty] \mid (\forall x' \in E)(d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon)\}$$

et

$$\delta_{x,\varepsilon} = \sup C_{x,\varepsilon}. \tag{4}$$

Comme f est continue sur E, on a $C_{x,\varepsilon} \neq \varnothing$ et donc $\delta_{x,\varepsilon} > 0$. Montrons que $\delta_{x,\varepsilon} \in C_{x,\varepsilon}$, de sorte que $\delta_{x,\varepsilon}$ est le plus grand élément de $C_{x,\varepsilon}$. En effet, soit $x' \in E$ et $d_X(x, x') < \delta_{x,\varepsilon}$. En vertu de (4), $d_X(x, x') < \delta$ pour un $\delta \in C_{x,\varepsilon}$ et, par conséquent, $d_Y(f(x), f(x')) < \varepsilon$. Il est évident que f est uniformément continue sur E si et seulement si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ appartenant à $C_{x,\varepsilon}$ quel que soit $x \in E$ ou, ce qui revient au même, si $\inf\limits_{x \in E} \delta_{x,\varepsilon} > 0$ pour tout $\varepsilon > 0$. Mais il peut arriver que $\inf\limits_{x \in E} \delta_{x,\varepsilon} = 0$ pour un (ou même pour tout) $\varepsilon > 0$.

EXEMPLE 4. Les notations étant celles de la définition 7, soient $X = Y = \mathbf{R}$, $f(x) = x^2$, $E = [0, +\infty[$ et $\varepsilon > 0$. Si $x > 0$ et $x' = \sqrt{x^2 + \varepsilon}$, on a $|f(x') - f(x)| = \varepsilon$ et donc $\delta_{x,\varepsilon} = \sqrt{x^2 + \varepsilon} - x$ (fig. 8). Quel que soit $\varepsilon > 0$,

$$\sqrt{x^2 + \varepsilon} - x = \frac{\varepsilon}{\sqrt{x^2 + \varepsilon} + x} \to 0 \text{ quand } x = +\infty.$$

C'est pourquoi, $\inf\limits_{x \in E} \delta_{x,\varepsilon} = 0$ pour tout $\varepsilon > 0$ et, par conséquent, bien que la fonction x^2 soit continue sur l'intervalle $[0, +\infty[$, sa continuité n'est pas

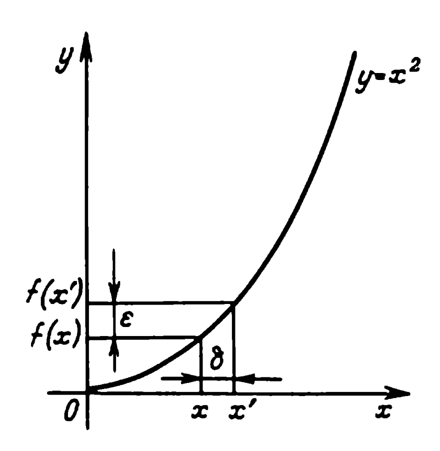


Fig. 8

uniforme. D'autre part, si x et x' appartiennent à l'intervalle $[0, a[$ où a est un nombre strictement positif arbitraire,

$$|f(x') - f(x)| = (x + x')|x' - x| \leqslant 2a|x' - x|,$$

de sorte que $|x' - x| < \varepsilon/(2a) \Rightarrow |f(x') - f(x)| < \varepsilon$ quel que soit $\varepsilon > 0$. C'est pourquoi, $\delta_{x,\varepsilon} \geqslant \varepsilon/(2a)$ pour tout $x \in [0, a[$ et, donc, la fonction x^2 est uniformément continue sur l'intervalle $[0, a[$. (Ainsi donc, la fonction continue peut être uniformément continue sur un ensemble et ne pas être uniformément continue sur l'autre.)

Comment expliquer que la fonction $f(x) = x^2$ n'est pas uniformément continue sur l'intervalle $[0, +\infty[$ et est uniformément continue sur $[0, a[$ $(0 < a < +\infty)$? Peut-être par le fait que le premier intervalle n'est pas borné et les intervalles $[0, a[$ le sont ? Ou bien par le fait que la fonction n'est pas bornée sur $[0, +\infty[$ tout en l'étant sur $[0, a[$? Les exemples qui suivent montrent que l'explication est tout à fait autre.

EXEMPLE 5. La fonction $f(x) = x$ est uniformément continue sur **R** (et, par conséquent, sur $[0, +\infty[$), bien que **R** ne soit pas borné et f ne soit pas bornée sur **R**. En effet, quel que soit $\varepsilon > 0$, $|x' - x| < \varepsilon \Leftrightarrow |f(x') - f(x)| < \varepsilon$, de sorte que $\delta_{x,\varepsilon} = \varepsilon$ pour tout $x \in \mathbf{R}$.

EXEMPLE 6. La fonction $f(x) = \sin x^2$ est bornée et continue, mais sur l'intervalle $[0, +\infty[$ elle n'est pas uniformément continue. En effet, si $x = \sqrt{\pi n}$ et $x' = \sqrt{\pi n + \pi/2}$ où $n \in \mathbf{N}$, on a $f(x) = 0$ et $f(x') = \pm 1$, de sorte que quel que soit $\varepsilon \in]0, 1]$,

$$\delta_{\sqrt{\pi n},\varepsilon} < \sqrt{\pi n + \frac{\pi}{2}} - \sqrt{\pi n} = \frac{\pi/2}{\sqrt{\pi n + \pi/2} + \sqrt{\pi n}}$$

et par conséquent, $(0 \leqslant) \inf\limits_{x \geqslant \delta} \delta_{x,\varepsilon} \leqslant \inf\limits_{n \in \mathbf{N}} \delta_{\sqrt{\pi n},\varepsilon} = 0$.

EXEMPLE 7. La fonction $f(x) = \sin(1/x)$ est bornée et continue, l'intervalle $]0, 1]$ est borné, mais f n'est pas uniformément continue sur $]0, 1]$. En effet, soit $\varepsilon \in]0, 1]$. Pour tout $\delta > 0$ il existe un $n \in \mathbf{N}$ tel que $\dfrac{2}{\pi n(n + 1)} < \delta$. Supposons que $x = \dfrac{2}{n\pi}$, $x' = \dfrac{2}{(n + 1)\pi}$. Alors, $x, x' \in]0, 1]$ et $|x - x'| < \delta$, mais

$$|f(x) - f(x')| = \left|\sin\frac{n\pi}{2} - \sin\frac{(n + 1)\pi}{2}\right| = 1 \geqslant \varepsilon.$$

EXERCICE 5. Montrer que toute fonction uniformément continue sur un intervalle borné est bornée (de sorte qu'une fonction non bornée sur un intervalle borné n'est pas uniformément continue; telle est, par exemple, la fonction $f(x) = 1/x$ sur l'intervalle $]0, 1]$).

Dans les exemples cités il s'agissait des intervalles sur lesquels on avait, parmi les fonctions continues, des fonctions uniformément et non uniformément continues. Cependant, il existe une classe d'intervalles sur lesquels toute fonction réelle continue est uniformément continue. Cette classe est celle des segments, ce qui s'explique par le fait que ces intervalles sont des compacts. Quant à la propriété des fonctions, elle se généralise à toutes les applications continues des compacts dans les espaces métriques.

THÉORÈME 14. *Une application f d'un espace métrique X dans un espace métrique Y est uniformément continue sur tout compact $K \subset D_f$ où elle est continue.*

Démonstration. Les boules ouvertes $U(x\,;\ \delta_{x,\varepsilon/2}/2)$ $(x \in K)$ *) recouvrent K. Comme K est un compact, il existe un sous-recouvrement fini de ces boules, i.e. il existe des points $x_1, \ldots, x_n \in K$ tels que

$$K \subset \bigcup_{k=1}^{n} (x_k\,;\ \delta_{x_k,\varepsilon/2}/2). \tag{5}$$

Posons

$$\delta = \min_{k \in [1,n]} \delta_{x_k,\varepsilon/2}/2$$

(de sorte que $\delta > 0$). Alors

$$(\forall x \in K)(\forall x' \in K)(d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon),$$

si bien que f est uniformément continue sur K. En effet, soient $x, x' \in K$ et $d_X(x, x') < \delta$. Il existe en vertu de (5) un $k \in [1, n]$ tel que $x \in U(x_k\,;\ \delta_{x_k,\varepsilon/2}/2)$, i.e. $d_X(x, x_k) < \delta_{x_k,\varepsilon/2}/2$. Vu que $d_X(x, x') < \delta \leqslant \delta_{x_k,\varepsilon/2}/2$, on obtient

$$d_X(x', x_k) \leqslant d_X(x', x) + d_X(x, x_k) < 2\delta_{x_k,\varepsilon/2}/2 = \delta_{x_k,\varepsilon/2}.$$

Puisque $d_X(x, x_k) < \delta_{x_k,\varepsilon/2}$ et $\delta_{x_k,\varepsilon/2} \in C_{x_k,\varepsilon/2}$, on en conclut que

$$d_Y(f(x), f(x')) \leqslant d_Y(f(x), f(x_k)) + d_Y(f(x'), f(x_k)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui démontre le théorème.

Citons encore une démonstration du théorème 14. D'après le théorème 9, la fonction semi-continue inférieurement sur un ensemble compact possède la plus petite valeur. Il s'ensuit que pour démontrer le théorème 14, il suffit de démontrer que $\delta_{x,\varepsilon}$ est semi-continue inférieurement pour tout $\varepsilon > 0$ donné, ce qui permet de conclure que sa plus petite valeur sur K appartient à $C_{x,\varepsilon}$ pour tout $x \in K$. Soit donc $x \in K$ et $\delta < \delta_{x,\varepsilon}$ pour un $\varepsilon > 0$ donné. Prenons des nombres δ', δ'' satisfaisant aux doubles inégalités $\delta < \delta'' < \delta' < \delta_{x,\varepsilon}$ et soit

$$\varepsilon' = \sup_{x' \in K \cap B(x;\ \delta')} d_Y(f(x), f(x')). \tag{6}$$

Comme $B(x;\ \delta')$ est fermé dans X (voir n° 1.4.5), $K \cap B(x;\ \delta')$ est fermé dans K (voir théorème 1.4.5) et donc, est un compact en tant qu'un sous-ensemble fermé d'un compact (voir théorème 5). Il en résulte que ε' est la plus grande valeur de la fonction continue $d_Y(f(x'), f(x))$ de x' sur $K \cap B(x;\ \delta')$ (voir n° 5, remarque à la propriété 1°). Or

$$x' \in K \cap B(x;\ \delta') \Rightarrow d_X(x, x') < \delta_{x,\varepsilon} \Rightarrow d_Y(f(x'), f(x)) < \varepsilon.$$

Par conséquent, on a en particulier $\varepsilon' < \varepsilon$. C'est pourquoi il existe un nombre strictement positif η qui satisfait aux conditions $\eta \leqslant \delta' - \delta''$ et

$$x_1 \in K \cap U(x;\ \eta) \Rightarrow d_Y(f(x_1), f(x)) < \varepsilon - \varepsilon'. \tag{7}$$

---

*) Si $\delta_{x,\varepsilon} = +\infty$, on entend évidemment par $U(x;\ \delta_{x,\varepsilon/2}/2)$ tout le X.

Si $x_1 \in K \cap U(x;\eta)$ et $x' \in K \cap U(x_1;\delta'')$, on obtient $d_X(x',x) \leqslant d_X(x,x_1) + d_X(x_1,x') < < \eta + \delta'' \leqslant \delta'$ et, en vertu de (6) et (7),

$$d_Y(f(x'),f(x_1)) \leqslant d_Y(f(x'),f(x)) + d_Y(f(x),f(x_1)) < \varepsilon' + (\varepsilon - \varepsilon') = \varepsilon.$$

Ainsi donc, si $x_1 \in K \cap U(x;\eta)$, il vient

$$(\forall x' \in K)(d_X(x',x_1) < \delta'' \Rightarrow d_Y(f(x'),f(x_1)) < \varepsilon),$$

i.e. $\delta'' \in C_{x,\varepsilon}$ et, par conséquent, $\delta_{x,\varepsilon} \geqslant \delta'' > \delta$. Ceci démontre que pour tout $\delta < \delta_{x,\varepsilon}$ il existe un $\eta > 0$ tel que $\delta_{x_1,\varepsilon} > \delta$ pour tout $x_1 \in K \cap U(x;\eta)$. Mais cela signifie que $\delta_{x,\varepsilon}$ est semi-continue inférieurement sur K.

EXERCICE 6. Montrer que $\delta_{x,\varepsilon}$ n'est pas en général continue sur K.

EXERCICE 7. Les notations étant celles de la définition 7, supposons que f soit continue sur E et que

$$B_{x,\varepsilon} = \{\delta \in]0,+\infty] \mid x',x'' \in U(x;\delta) \cap E \Rightarrow d_Y(f(x'),f(x'')) < \varepsilon\}$$

et $\gamma_{x,\varepsilon} = \sup B_{x,\varepsilon}$. Etant donné un $\varepsilon > 0$, $\gamma_{x,\varepsilon}$ est soit fini partout sur E, soit infini partout sur E. Montrer que la fonction $\gamma_{x,\varepsilon}$ est continue sur E pour chaque $\varepsilon > 0$ pour lequel elle prend des valeurs finies et utiliser ce résultat pour la démonstration du théorème 14. (Notice. Montrer que si $\delta_{x,\varepsilon}$ prend des valeurs finies, $|\delta_{x',\varepsilon} - \delta_{x'',\varepsilon}| \leqslant \eta$ pour tous $x',x'' \in E$ tels que $d_X(x',x'') < \eta$, de sorte que la fonction $\delta_{x,\varepsilon}$ est même uniformément continue sur E.)

COROLLAIRE. *Toute fonction réelle continue sur un segment est uniformément continue.*

EXERCICE 8. Montrer que sur tout intervalle non compact il existe une fonction continue qui n'est pas uniformément continue.

§ 3. Frontière d'un ensemble *)

1. Notion de frontière

DÉFINITION 1. Soient E un sous-ensemble d'un espace topologique X et $\complement_X E$ le complémentaire de E dans X. Un point $x \in X$ est appelé *point frontière de E* si tout voisinage de x a une intersection non vide avec E et avec $\complement_X E$, i.e. x est un point adhérent à E et à $\complement_X E$. L'ensemble de tous les points frontières $x \in X$ de E est appelé *frontière de l'ensemble E* dans X ; nous la noterons Fr E. Ainsi donc,

(1) $$\text{Fr}\, E = [E] \cap [\complement_X E].$$

Comme $[\complement_X E] = \complement_X]E[$ en vertu de la propriété 1.4.4, 2°, d), on a Fr $E = [E] \cap \complement_X]E[= [E] \setminus]E[$, i.e. *la frontière d'un ensemble est la différence entre son adhérence et son intérieur.*

D'une façon plus générale, un point $x \in X$ peut se trouver dans une des relations suivantes avec l'ensemble $E \subset X$:

1) E est un voisinage du point x, i.e. x est un *point intérieur à l'ensemble E.*

*) Nous aurons besoin de ce paragraphe dans le chapitre 11.

2) $C_X E$ est un voisinage du point x, i.e. *x est un point intérieur à l'ensemble* $C_X E$; on dit alors que x est un *point extérieur à E.*

3) Ni E ni $C_X E$ ne sont des voisinages du point x. Alors x est un *point frontière de E.* En effet, soit U un voisinage arbitraire du point x. Le voisinage U n'est pas contenu dans E car autrement E serait aussi un voisinage de x ; par conséquent, U a une intersection non vide avec $C_X E$. Pour les mêmes raisons, U n'est pas contenu dans $C_X E$ et, par conséquent, a une intersection non vide avec E.

EXEMPLE 1. Soient $X = \mathbf{R}$, $E = \langle \alpha, \beta \rangle$ un intervalle de $\mathbf{R}$ où $\alpha, \beta \in \mathbf{R}$ et $\alpha < \beta$ (fig. 9). Si $\alpha < x < \beta$, x est un point intérieur à $\langle \alpha, \beta \rangle$ dans $\mathbf{R}$ parce que cet intervalle est un voisinage de x dans $\mathbf{R}$. Si $x < \alpha$ ou $x > \beta$, x est un point extérieur à $\langle \alpha, \beta \rangle$. Enfin, si $x = \alpha$ ou $x = \beta$, x est un point frontière de $\langle \alpha, \beta \rangle$.

EXEMPLE 2. Soient $X = \mathbf{R}^n$, $E = B(a\,;\, r)$ (fig. 10). Montrons que

$$\|x - a\| < r \Leftrightarrow x \in \,]B(a\,;\, r)[,$$
$$\|x - a\| > r \Leftrightarrow x \in \,]C_{\mathbf{R}^n} B(a\,;\, r)[,$$
$$\|x - a\| = r \Leftrightarrow x \in \operatorname{Fr} B(a\,;\, r).$$

a) Soit $\|x - a\| < r$, i.e. $x \in U(a\,;\, r)$. Comme $U(a\,;\, r)$ est un ensemble ouvert et $U(a\,;\, r) \subset B(a\,;\, r)$, on a $U(a\,;\, r) \subset \,]B(a\,;\, r)[$. Donc, $\|x - a\| < r \Rightarrow x \in \,]B(a\,;\, r)[$.

b) Soit $\|x - a\| > r$. Supposons que $\|x - a\| - r = \varrho$, de sorte que $\varrho > 0$, et soit $y \in U(x\,;\, \varrho)$, i.e. $\|y - x\| < \|x - a\| - r$. Alors $\|y - a\| \geqslant$ $\geqslant \|x - a\| - \|y - x\| > r$, i.e. $y \notin B(a\,;\, r)$. Ainsi donc, $U(x\,;\, \varrho) \cap$ $\cap B(a\,;\, r) = \varnothing$, i.e. $U(x\,;\, \varrho) \subset C_{\mathbf{R}^n} B(a\,;\, r)$ et, par conséquent, $x \in \,]C_{\mathbf{R}^n} B(a\,;\, r)[$.

c) Enfin, soit $\|x - a\| = r$. Considérons une demi-droite issue de a et passant par x, i.e. l'ensemble

$$\{x_t = (1 - t)a + tx \mid t \geqslant 0\}.$$

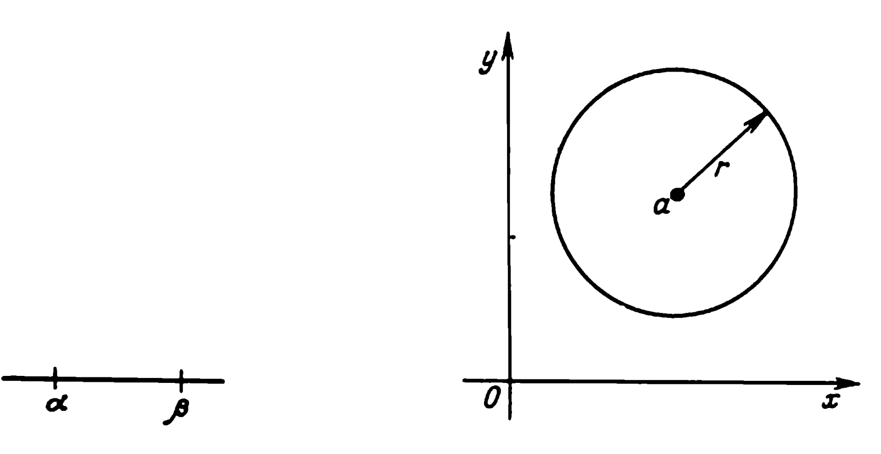


Fig. 9 Fig. 10

Comme $\|x_t - a\| = \|t(x-a)\| = tr$, x_t est un point intérieur à $B(a\,;\,r)$ si $t < 1$, et un point extérieur si $t > 1$. Mais tout voisinage du point $x = x_1$ contient les uns et les autres points x_t. En effet, puisque

$$\|x_t - x\| = \|(t-1)(x-a)\| = |t-1|\,\|x-a\| = |t-1|r,$$

on a $|t-1| < \dfrac{\varepsilon}{r} \Rightarrow \|x_t - x\| < \varepsilon$, et t peut être soit inférieur, soit supérieur à 1.

d) Puisque tout $x \in \mathbf{R}^n$ vérifie une seule des relations $\|x-a\| < r$, $\|x-a\| > r$, $\|x-a\| = r$ et que les relations

$$x \in \,]B(a\,;\,r)[,\ x \in \,]C_{\mathbf{R}^n}B(a\,;\,r)[,\ x \in \operatorname{Fr} B(a\,;\,r)$$

sont incompatibles deux à deux, il vient

$$\begin{aligned}]B(a\,;\,r)[&= \{x \in \mathbf{R}^n \mid \|x-a\| < r\} = U(a\,;\,r),\\ ]C_{\mathbf{R}^n}B(a\,;\,r)[&= \{x \in \mathbf{R}^n \mid \|x-a\| > r\},\\ \operatorname{Fr} B(a\,;\,r) &= \{x \in \mathbf{R}^n \mid \|x-a\| = r\} = S(a\,;\,r). \end{aligned}$$

Il résulte directement de la formule (1) que *la frontière de tout ensemble* (en tant qu'une intersection d'ensembles fermés) *est fermée* et que

(2) $$\operatorname{Fr} E = \operatorname{Fr}(C_X E).$$

Il est évident que $\operatorname{Fr} \varnothing = \operatorname{Fr} X = \varnothing$.

EXERCICE 1. Montrer qu'un ensemble est fermé si et seulement s'il contient sa frontière.

EXERCICE 2. Montrer qu'un espace X est connexe si et seulement si $\varnothing$ et X sont les seules parties de X ayant une frontière vide.

Le théorème suivant est très important pour la théorie des figures quarrables.

THÉORÈME 1. *La frontière de la réunion, de l'intersection et de la différence de deux ensembles est contenue dans la réunion de leurs frontières* (fig. 11).

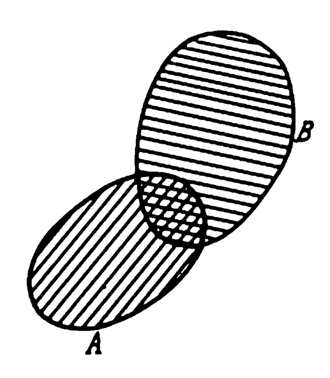


Fig. 11

Démonstration. Rappelons que $[A \cup B] = [A] \cup [B]$ (voir propriété 1.4.4, 6°). 1) En vertu de (1),

$$\begin{aligned}\mathrm{Fr}\,(A \cup B) &= [A \cup B] \cap [\mathrm{C}_X(A \cup B)] = \\ &= ([A] \cup [B]) \cap [\mathrm{C}_X A \cap \mathrm{C}_X B] \subset ([A] \cup [B]) \cap [\mathrm{C}_X A] \cap [\mathrm{C}_X B] = \\ &= ([A] \cap [\mathrm{C}_X A] \cap [\mathrm{C}_X B]) \cup ([B] \cap [\mathrm{C}_X A] \cap [\mathrm{C}_X B]) \subset \\ &\subset ([A] \cap [\mathrm{C}_X A]) \cup ([B] \cap [\mathrm{C}_X B]) = \mathrm{Fr}\,A \cup \mathrm{Fr}\,B.\end{aligned}$$

2) En vertu de (2) et 1)

$$\begin{aligned}\mathrm{Fr}\,(A \cap B) = \mathrm{Fr}\,(\mathrm{C}_X(A \cap B)) &= \mathrm{Fr}\,(\mathrm{C}_X A \cup \mathrm{C}_X B) \subset \\ &\subset \mathrm{Fr}\,(\mathrm{C}_X A) \cup \mathrm{Fr}\,(\mathrm{C}_X B) = \mathrm{Fr}\,A \cup \mathrm{Fr}\,B.\end{aligned}$$

3) En vertu de (2) et 2)

$$\mathrm{Fr}\,(A \setminus B) = \mathrm{Fr}\,(A \cap \mathrm{C}_X B) \subset \mathrm{Fr}\,A \cup \mathrm{Fr}\,(\mathrm{C}_X B) = \mathrm{Fr}\,A \cup \mathrm{Fr}\,B.$$

2. Frontière de l'ensemble dans $\mathbf{R}^n$

Rappelons qu'au n° 2 du § 1 il a été établi que tout segment de $\mathbf{R}^n$ est connexe.

THÉORÈME 2. *Soit E un ensemble de $\mathbf{R}^n$. Si $a \in E$ et $b \in \mathrm{C}_{\mathbf{R}^n} E$* ($= \mathbf{R}^n \setminus E$), *le segment $[a, b]$ contient au moins un point frontière de E.*

Démonstration. Posons $A = [a, b] \cap E$ et $B = [a, b] \cap \mathrm{C}_{\mathbf{R}^n} E$. On a $a \in A$ et $b \in B$, de sorte que A et B ne sont pas vides. Comme $[a, b] = [a, b] \cap (E \cup \mathrm{C}_{\mathbf{R}^n} E) = A \cup B$ et $A \cap B = \varnothing$, le segment $[a, b]$ contient, en vertu de sa connexité, un point c appartenant à $A \cap [B]$ ou à $[A] \cap B$. Dans les deux cas, $c \in [A] \cap [B]$. Mais $A \subset E$ et $B \subset \mathrm{C}_{\mathbf{R}^n} E$ entraînent $[A] \cap [B] \subset [E] \cap [\mathrm{C}_{\mathbf{R}^n} E]$. Compte tenu de la formule (1), on en conclut que $c \in \mathrm{Fr}\,E$.

COROLLAIRE. *$\varnothing$ et $\mathbf{R}^n$ sont les seuls ensembles de $\mathbf{R}^n$ possédant une frontière vide.*

En effet, si $E \subset \mathbf{R}^n$, $E \neq \varnothing$ et $E \neq \mathbf{R}^n$, il existe au moins un point $a \in E$ et au moins un point $b \in \mathrm{C}_{\mathbf{R}^n} E$ tels que $[a, b] \cap \mathrm{Fr}\,E \neq \varnothing$ en vertu du théorème 1. Il en résulte à plus forte raison que $\mathrm{Fr}\,E \neq \varnothing$.

THÉORÈME 3. *La frontière de tout ensemble borné de $\mathbf{R}^n$ est un compact, qui n'est pas vide si l'ensemble n'est pas vide.*

Démonstration. Soit E un ensemble borné de $\mathbf{R}^n$, i.e. il existe un $r > 0$ tel que $E \subset B(0\,;\, r)$. Comme la boule $B(0\,;\, r)$ est fermée (voir n° 1.4.5), $[E] \subset B(0\,;\, r)$. En vertu de la formule (1), $\mathrm{Fr}\,E \subset [E]$. Par conséquent, $\mathrm{Fr}\,E \subset B(0\,;\, r)$, si bien que $\mathrm{Fr}\,E$ est bornée. D'autre part, on a déjà vu au n° 1 que $\mathrm{Fr}\,E$ est fermée. D'après le théorème 13 du § 2, tout ensemble fermé et borné de $\mathbf{R}^n$ est un compact. Enfin, étant borné, $E \neq \mathbf{R}^n$. Donc, si $E \neq \varnothing$, le corollaire du théorème 2 dit que $\mathrm{Fr}\,E \neq \varnothing$.

DÉFINITION 2. On appelle *ε-voisinage d'un ensemble E* dans un espace métrique X la réunion $U(E\,;\, \varepsilon)$ des ε-voisinages de tous les points de cet

ensemble :

$$U(E\,;\,\varepsilon) = \bigcup_{x \in E} U(x\,;\,\varepsilon).$$

EXERCICE 3. Montrer que $U(E\,;\,\varepsilon)$ est un ensemble de tous les points de X dont la distance à E est strictement inférieure à ε.

THÉORÈME 4. *Soient* $E, F \subset \mathbf{R}^n$. *Si* $E \subset F$ *et* $\delta = \varrho(E, \operatorname{Fr} F) > 0$, *on a* $U(E\,;\,\delta) \subset F$.

Démonstration. Rappelons (voir n° 8 du § 2) que $\varrho(A, B)$, où A et B sont des sous-ensembles d'un espace métrique muni de la métrique d, désigne la distance entre eux :

$$\varrho(A, B) = \inf\{d(a, b) \mid a \in A,\ b \in B\}.$$

Soit $x \in U(E\,;\,\delta)$. Il existe alors dans E un point x' tel que $x \in U(x'\,;\,\delta)$, i.e. $d(x', x) < \delta$. Si x n'appartenait pas à F, il existerait en vertu du théorème 2 un point $x_0 \in [x', x] \cap \operatorname{Fr} F$ puisque $x' \in F$. Il s'ensuit que $x_0 = x' + \lambda(x - x')$ où $\lambda \in [0, 1]$, ce qui implique

$$d(x', x_0) = \|x_0 - x'\| = \|\lambda(x - x')\| = \lambda\|x - x'\| \leqslant \|x - x'\| = d(x', x),$$

de sorte qu'on obtiendrait $\delta = \varrho(E, \operatorname{Fr} F) \leqslant d(x', x_0) \leqslant d(x', x) < \delta$.

CHAPITRE 4

COMPLÉTUDE ET THÉORÈMES DU POINT FIXE

§ 1. Espace métrique complet

1. Suite de Cauchy. Complétude

La notion de suite de Cauchy dans **R** se généralise aisément à tous les espaces métriques.

DÉFINITION 1. On dit que la suite (x_n) d'éléments d'un espace métrique muni de la métrique d est une *suite de Cauchy* si pour tout nombre $\varepsilon > 0$ il existe un nombre N tel que

$$d(x_p, x_q) < \varepsilon \qquad \text{pour tous } p, q \geqslant N,$$

ou, ce qui est équivalent,

$$d(x_{n+m}, x_n) < \varepsilon \qquad \text{pour tout } n \geqslant N \text{ et tout } m \in \mathbf{N}.$$

EXERCICE 1. Il résulte de la convergence d'une suite la convergence vers la même limite de ses sous-suites. Montrer que pour une suite de Cauchy la réciproque est aussi vraie : la convergence de l'une quelconque de ses sous-suites entraîne la convergence de cette suite.

Toute suite convergente est une suite de Cauchy. En effet, si $x_n \to x$, il existe pour tout $\varepsilon > 0$ un nombre N tel que $d(x_n, x) < \varepsilon/2$ quel que soit $n \geqslant N$, ce qui implique $d(x_p, x_q) \leqslant d(x_p, x) + d(x_q, x) < \varepsilon$ pour tous $p, q \geqslant N$.

La réciproque n'est en général pas vraie. En effet, supposons que **Q** (ensemble de tous les nombres rationnels) soit muni d'une métrique ordinaire $d(x, y) = |x - y|$ induite par celle de **R** et que (r_n) soit une suite de **Q** qui tend dans **R** vers un nombre irrationnel α (par exemple, la suite des approximations décimales par défaut de $\sqrt{2}$). D'après ce qui vient d'être démontré, (r_n) est une suite de Cauchy dans **R** et, par conséquent, dans **Q** aussi. Mais elle n'est pas convergente dans **Q** car autrement elle aurait dans **R** deux limites distinctes.

DÉFINITION 2. Un espace métrique est dit *complet* si toute suite de Cauchy y converge. Un espace normé est dit complet s'il est complet en tant qu'espace métrique muni de la métrique engendrée par sa norme.

EXEMPLE 1. On démontre que toute suite de Cauchy converge dans **R**, ce qui signifie que **R** est complet pour la métrique ordinaire.

EXERCICE 2. Montrer que **R** muni de la métrique $d(x, y) = |\text{Arctg}\, x - \text{Arctg}\, y|$ (voir n° 1.2.1) n'est pas complet, quoique cette métrique engendre la même convergence et la même topologie que la métrique ordinaire.

EXEMPLE 2. *Quel que soit n, l'espace $\mathbf{R}^n$ est complet pour la métrique euclidienne.* En effet, soit $(x^{(p)})$ une suite de Cauchy dans l'espace euclidien n-dimensionnel $\mathbf{R}^n$. Pour tout $\varepsilon > 0$ il existe alors un nombre N tel que

$$\sqrt{\sum_{i=1}^{n} (x_i^{(p)} - x_i^{(q)})^2} < \varepsilon \text{ quels que soient } p,\ q \geqslant N.$$

Comme $|x_k^{(p)} - x_k^{(q)}| \leqslant \sqrt{\sum_{i=1}^{n} (x_i^{(p)} - x_i^{(q)})^2}$ $(k = 1, 2, \ldots, n)$, on a à plus forte raison $|x_k^{(p)} - x_k^{(q)}| < \varepsilon$ pour tous p, $q \geqslant N$. Donc, toutes les suites $(x_k^{(p)})$ $(k \in [1, n])$ sont des suites de Cauchy dans $\mathbf{R}$ et, par suite, elles convergent dans $\mathbf{R}$ car cet espace est complet. Soient $x_k = \lim x_k^{(p)}$ $(k \in [1, n])$ et $x = (x_1, \ldots, x_n)$. Comme la convergence dans $\mathbf{R}$ est une convergence en coordonnées (voir exemple 1.2.1 et n° 1.4.3), $x^{(p)} \to x$. Ainsi donc, toute suite de Cauchy converge dans $\mathbf{R}^n$.

EXEMPLE 3. *L'espace hilbertien l_2* (voir exemple 1.2.2) *est complet.* Rappelons que l_2 est un espace réel normé ; il est formé de toutes les suites $x = (x_n)$ de nombres réels pour lesquelles la série $\sum x_n^2$ converge ; la norme sur l_2 est définie par $\|x\| = \sqrt{\sum_{n=1}^{\infty} x_n^2}$. Soit $(x^{(p)})$ une suite de Cauchy dans l_2, de sorte que pour tout $\varepsilon > 0$ il existe un nombre N tel que

$$\sqrt{\sum_{k=1}^{\infty} (x_k^{(p)} - x_k^{(q)})^2} < \varepsilon/2 \text{ quels que soient } p,\ q \geqslant N.$$

Comme $|x_k^{(p)} - x_k^{(q)}| \leqslant \sqrt{\sum_{i=1}^{n} (x_i^{(p)} - x_i^{(q)})^2}$ pour tout $n \in \mathbf{N}$, $(x_n^{(p)})$ est une suite de Cauchy dans $\mathbf{R}$ pour chaque n fixé et, par conséquent, c'est une suite convergente. Soient $x_n = \lim x_n^{(p)}$ et $x = (x_n)$.

Fixons un $q \geqslant N$ et passons dans les inégalités

$$\sqrt{\sum_{k=1}^{m} (x_k^{(p)} - x_k^{(q)})^2} < \varepsilon/2 \quad (m \in \mathbf{N})$$

à la limite quand $p \to +\infty$ et puis quand $m \to +\infty$. On obtient que la série $\sum_{k=1}^{\infty} (x_k^{(p)} - x_k^{(q)})^2$ est convergente, de sorte que $x - x^{(q)} \in l_2$ et

$$\text{(1)} \qquad \sqrt{\sum_{k=1}^{\infty} (x_k^{(p)} - x_k^{(q)})^2} \leqslant \frac{\varepsilon}{2} < \varepsilon \text{ pour tout } q \geqslant N.$$

Puisque $x^{(q)} \in l_2$ et $x - x^{(q)} \in l_2$, on a $x = (x - x^{(q)}) + x^{(q)} \in l_2$. La relation (1) signifie alors que $x^{(q)} \to x$ dans l_2 quand $q \to +\infty$, de sorte que toute suite de Cauchy dans l_2 est convergente.

EXERCICE 3. Soit m un espace normé de toutes les suites bornées $x = (x_n)$ de nombres réels dont la norme est $\|x\| = \sup_n |x_n|$. Démontrer que l'espace m est complet.

EXERCICE 4. Montrer qu'un espace métrique est complet si et seulement si toute suite strictement décroissante de boules fermées dont les rayons tendent vers 0 a une intersection non vide.

2. Espaces d'applications continues bornées

Rappelons qu'un ensemble de l'espace métrique est dit borné s'il existe une boule qui le contient (voir définition 3.2.3). Une application dans un espace métrique est dite *bornée* si l'ensemble de ses valeurs est borné. Dans ce numéro, X et Y désignent des espaces métriques non vides munis de métriques d_X et d_Y. Notons $C(X, Y)$ l'ensemble de toutes les applications continues bornées de X dans Y. Il n'est pas vide puisqu'il contient des applications constantes. Pour chaque couple (f, g) de ses éléments posons

$$d(f, g) = \sup_{x \in X} d_Y(f(x), g(x)).$$

Montrons que d est une *métrique* sur $C(X, Y)$:

a) soit y_0 un élément arbitraire de l'espace Y. Comme $f(X)$ et $g(X)$ sont bornés, il existe en vertu du n° 3.2.5 des nombres A et B tels que $d_Y(f(x), y_0) < A$ et $d_Y(g(x), y_0) < B$ pour tout $x \in X$. On a alors $d_Y(f(x), g(x)) < A + B$ pour tout $x \in X$ et, par conséquent, $d(f, g)$ est un nombre ;

b) soient $f, g, h \in C(X, Y)$. Comme $d_Y(f(x), g(x)) \leqslant d_Y(f(x), h(x)) + d_Y(g(x), h(x)) \leqslant d(f, h) + d(g, h)$ pour tout $x \in X$, on obtient $d(f, g) \leqslant d(f, h) + d(g, h)$;

c) $d(f, g) = 0 \Leftrightarrow d_Y(f(x), g(x)) = 0$ pour tout $x \in X \Leftrightarrow f(x) = g(x)$ pour tout $x \in X \Leftrightarrow f = g$.

La convergence dans $C(X, Y)$ est une convergence uniforme sur X, i.e. *$f_n \to f$ dans $C(X, Y)$ si et seulement si pour tout $\varepsilon > 0$ il existe un nombre N tel que $d_Y(f_n(x), f(x)) < \varepsilon$ pour tout $n > N$ et tout $x \in X$.* En effet, $f_n \to f$ dans $C(X, Y)$ signifie que $d(f_n, f) = \sup_{x \in X} d_Y(f_n(x), f(x)) \to 0$, i.e. il existe pour tout $\varepsilon > 0$ un nombre N tel que $\sup_{x \in X} d_Y(f_n(x), f(x)) < \varepsilon$ quel que soit $n > N$; mais ceci est équivalent à ce que $d_Y(f_n(x), f(x)) < \varepsilon$ pour tout $n > N$ et tout $x \in X$.

LEMME. *La limite f d'une suite uniformément convergente (f_n) d'applications continues de X dans Y est continue.*

Démonstration. Vu que (f_n) converge uniformément vers f sur X, il existe pour tout nombre $\varepsilon > 0$ un $n \in \mathbf{N}$ tel que

$$(2) \qquad d_Y(f_n(x), f(x)) < \varepsilon/3 \text{ pour tout } x \in X.$$

Soit $x_0 \in X$. Puisque f_n est continue, il existe un $\delta > 0$ tel que

(3) $$d_Y(f_n(x), f_n(x_0)) < \varepsilon/3 \text{ pour tout } x \in U(x_0, \delta).$$

Et comme

$$d_Y(f(x), f(x_0)) \leqslant d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(x_0)) + d_Y(f_n(x_0), f(x_0)),$$

il découle de (2) et (3) que

$$d_Y(f(x), f(x_0)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \text{ pour tout } x \in U(x_0, \delta),$$

i.e. f est continue en un point (arbitraire) x_0 de l'espace X.

THÉORÈME 1. *Si Y est complet, il en est de même pour l'espace $C(X, Y)$ quel que soit X.*

Démonstration. Soit (f_n) une suite de Cauchy dans $C(X, Y)$, i.e. quel que soit $\varepsilon > 0$ il existe un nombre N tel que $d(f_p, f_q) < \varepsilon$ pour tous $p, q \geqslant N$. Comme $d_Y(f_p(x), f_q(x)) \leqslant d(f_p, f_q)$, on obtient

(4) $$d_Y(f_p(x), f_q(x)) < \varepsilon \text{ pour tous } p, q \geqslant N \text{ et } x \in X,$$

de sorte que $(f_n(x))$ est pour tout $x \in X$ une suite de Cauchy dans Y et, par conséquent, c'est une suite convergente en vertu de la complétude de Y. Soit $f(x) = \lim f_n(x)$ et soit un $y_0 \in Y$ quelconque. Comme f_N est bornée, il existe un $C > 0$ tel que $d_Y(f_N(x), y_0) < C$ pour tout $x \in X$. On obtient alors en vertu de (4)

$$d_Y(f_n(x), y_0) \leqslant d_Y(f_n(x), f_N(x)) + d_Y(f_N(x), y_0) < \varepsilon + C$$

pour tout $n \geqslant N$ et tout $x \in X$. En passant ici à la limite lorsque $n \to +\infty$, on obtient en vertu de la continuité de d_Y que $d_Y(f(x), y_0) \leqslant \varepsilon + C$ pour tout $x \in X$, donc f est bornée. D'autre part, le passage dans (4) à la limite lorsque $q \to +\infty$ donne

(5) $$d_Y(f_p(x), f(x)) \leqslant \varepsilon \text{ pour tous } p \geqslant N \text{ et } x \in X,$$

i.e. (f_p) converge uniformément vers f et, par conséquent, f est continue d'après le lemme. Ainsi donc, $f \in C(X, Y)$ et il résulte alors de (5) que $d(f_p, f) \leqslant \varepsilon$ pour tout $p \geqslant N$, i.e. $\lim f_p = f$ dans $C(X, Y)$.

COROLLAIRE. *L'espace $C(\Delta)$ de toutes les fonctions réelles continues sur un segment non dégénéré Δ* (voir exemple 1.2.3) *est complet.*

En effet, comme Δ est un compact, toutes les fonctions de $C(\Delta)$ sont bornées. Et puisque $C(\Delta)$ est muni de la norme $\|f\| = \max_{t \in \Delta} |f(t)|$, la métrique qu'elle engendre est

$$d(f, g) = \max_{t \in \Delta} |f(t) - g(t)| = \sup_{t \in \Delta} |f(t) - g(t)|.$$

Donc, $C(\Delta) = C(\Delta, \mathbf{R})$ et, comme $\mathbf{R}$ est complet, il en est de même pour $C(\Delta)$ d'après le théorème 1.

3. Sous-espace fermé d'un espace métrique complet

On appelle *sous-espace d'un espace métrique* X tout ensemble $E \subset X$ muni d'une métrique induite par celle de X, i.e. d'une métrique d_E définie pour tous $x, y \in E$ par la condition $d_E(x, y) = d_X(x, y)$. Il est évident que (x_n) est une suite de Cauchy dans E si et seulement si (x_n) est une suite de Cauchy dans X formée par les points $x_n \in E$, et $x_n \to x$ dans E si x_n et x appartiennent à E et $x_n \to x$ dans X.

Le sous-espace fermé E d'un espace métrique complet X est complet. En effet, soit (x_n) une suite de Cauchy dans E et, par conséquent, dans X. Puisque X est complet, (x_n) converge vers un $x \in X$. Mais comme E est fermé, $x \in E$ (voir propriété 1.4.5, 7°). Par conséquent, $x_n \to x$ dans E.

§ 2. Théorèmes du point fixe

1. Notion de point fixe. Théorème de Banach

DÉFINITION 1. Soient E un ensemble et A un opérateur sur E, i.e. une application de E dans E. On dit que $x \in E$ est un *point fixe de l'opérateur* A si $Ax = x$.

EXEMPLE 1. *Toute application continue d'un segment de droite dans lui-même possède un point fixe.* En effet, raisonnons par absurde et supposons qu'il existe une application f du segment $[a, b]$ de la droite numérique dans $[a, b]$, qui ne possède pas de point fixe. Considérons l'ensemble $E = \{t \in [a, b] | f(t) > t\}$. Comme f est continue, E est un ouvert dans $[a, b]$ (voir corollaire du théorème 2.2.3). D'autre part, vu que f n'a pas de point fixe, on peut écrire $E = \{t \in [a, b] | f(t) \geqslant t\}$, d'où il résulte que E est un fermé dans $[a, b]$ car f est continue. Puisque le segment $[a, b]$ est connexe, les seuls ensembles à la fois ouverts et fermés qu'il contient, sont le segment lui-même et l'ensemble vide. Cependant, $E \neq [a, b]$ puisque $f(b) < b$, et $E \neq \varnothing$ puisque $f(a) > a$.

EXERCICE 1. Montrer que toute application continue d'un segment $[a, b]$ de l'espace $\mathbf{R}^n$ dans ce segment possède un point fixe.

EXEMPLE 2. La rotation du cercle d'un angle $\alpha \in]0, 2\pi[$ autour de son centre est une application continue de ce cercle dans lui-même qui ne possède évidemment pas de point fixe.

Les théorèmes d'existence du point fixe de l'opérateur jouent un rôle important en analyse. L'un de ces théorèmes est le théorème de Banach démontré ci-après.

DÉFINITION 2. Etant donné l'espace métrique X muni de la métrique d, un opérateur A sur X est dit *contractant* s'il existe un nombre strictement positif $k < 1$ tel que

$$d(Ax, Ay) \leqslant kd(x, y) \text{ pour tous } x, y \in X.$$

Il est évident que *l'opérateur contractant est continu.*

THÉORÈME 1 (théorème du point fixe de Banach). *Tout opérateur contractant sur un espace métrique complet X possède un point fixe et un seul.*

Démonstration. Soit x_0 un point arbitraire de X. Pour tout $n \in \mathbf{N}$ posons $x_n = A^n x_0$ (i.e. $x_1 = Ax_0$, $x_2 = Ax_1 = A^2 x_0$, etc.). Montrons que

(x_n) est une suite de Cauchy dans X. En effet, quels que soient p, $q \in \mathbf{N}$,

$$d(x_p, x_q) = d(Ax_{p-1}, Ax_{q-1}) \leqslant kd(x_{p-1}, x_{q-1})$$

et donc pour tout $p \in \mathbf{N}$

$$d(x_{p+1}, x_p) \leqslant kd(x_p, x_{p-1}) \leqslant k^2 d(x_{p-1}, x_{p-2}) \leqslant \ldots \leqslant k^p d(x_1, x_0).$$

Il s'ensuit que pour tous m, $n \in \mathbf{N}$ on a les inégalités

$$d(x_{n+m}, x_n) \leqslant d(x_{n+1}, x_n) + d(x_{n+2}, x_{n+1}) + \ldots$$
$$\ldots + d(x_{n+m}, x_{n+m-1}) \leqslant (k^n + k^{n-1} + \ldots$$
$$\ldots + k^{n+m-1})d(x_1, x_0) < k^n \frac{d(x_1, x_0)}{1-k}.$$

Etant donné que $0 < k < 1$, on a $k^n \to 0$, d'où également $k^n \dfrac{d(x_1, x_0)}{1-k} \to 0$. Par conséquent, quel que soit $\varepsilon > 0$, il existe un nombre N tel que

$$d(x_{n+m}, x_n) < \varepsilon \text{ pour tout } n \geqslant N \text{ et tout } m \in \mathbf{N},$$

donc (x_n) est une suite de Cauchy. Comme X est complet, (x_n) est convergente ; soit $x = \lim x_n$. Puisque A est continu, $Ax = \lim Ax_n$. Or $Ax_n = x_{n+1} \to x$. Par conséquent, $Ax = x$, i.e. x est un point fixe de l'opérateur A. Enfin, si y est aussi un point fixe de A, $0 \leqslant d(x, y) = d(Ax, Ay) \leqslant kd(x, y)$, d'où $0 \leqslant (1-k)d(x, y) \leqslant 0$, i.e. $(1-k)d(x, y) = 0$. Comme $1 - k > 0$, on obtient $d(x, y) = 0$, i.e. $y = x$.

EXERCICE 2. Soient X un espace métrique complet et A une application d'une boule ouverte $U(x_0, r) \subset X$ dans X. Montrer que A possède un point fixe et un seul s'il existe un nombre $k > 0$ tel que $d(Ax, Ay) \leqslant kd(x, y)$ pour tous x, $y \in U(x_0, r)$ et $d(Ax_0, x_0) < r(1-k)$.

2. Renforcement du théorème de Banach

Le théorème 1 peut être renforcé (et utilisé avec succès sous cette forme dans la démonstration du théorème d'existence et d'unicité de la solution d'une équation différentielle) :

THÉORÈME 2. *Pour qu'un opérateur A sur un espace métrique complet possède un point fixe et un seul, il suffit que l'une de ses puissances A^p soit un opérateur contractant.*

Démonstration. Supposons que cette condition soit remplie. En vertu du théorème 1, l'opérateur A^p possède un point fixe x et un seul. On a alors $A^p(Ax) = A(A^p x) = Ax$, i.e. Ax est un point fixe de l'opérateur A^p et, par conséquent, $Ax = x$, i.e. x est aussi un point fixe de l'opérateur A. Si encore $Ay = y$, on obtient $A^p y = A^{p-1}(Ay) = A^{p-1}y = A^{p-2}(Ay) = A^{p-2}y = \ldots = Ay = y$. Comme x est le seul point fixe de l'opérateur A^p, on a $y = x$, i.e. x est encore le seul point fixe de l'opérateur A.

CHAPITRE 5

FONCTIONS DE PLUSIEURS VARIABLES

§ 1. Notions fondamentales

1. Fonction de $\mathbf{R}^n$ dans $\mathbf{R}$

La notion de fonction réelle de plusieurs variables repose sur le même fondement que celle de fonction d'une seule variable : c'est une dépendance entre les grandeurs ; mais cette fois il s'agit de la dépendance d'une grandeur par rapport à plusieurs autres grandeurs. Ainsi, le volume V du cylindre s'exprime par la formule $V = \pi R^2 H$ où R est le rayon de la base et H la hauteur, donc le volume est une fonction de deux variables : R et H. Celles-ci sont strictement positives et, pour le reste, arbitraires. La distance ϱ entre les points (x_1, y_1) et (x_2, y_2) d'une circonférence donnée par l'équation $(x - a)^2 + (y - b)^2 = r^2$ est exprimée par la formule

$$\varrho = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

donc ϱ est une « fonction de quatre variables » x_1, x_2, y_1, y_2 que vérifient les relations

$$(x_i - a)^2 + (y_i - b)^2 = r^2 \quad (i = 1, 2)$$

(et elles seules). Dans tous les exemples pareils, la valeur de la grandeur envisagée est définie par un ensemble de valeurs de plusieurs grandeurs, celles-ci pouvant vérifier certaines équations ou inéquations. Pour en déduire une notion analytique générale il faut :

1) faire abstraction du caractère concret des grandeurs définissant la grandeur donnée et envisager seulement les variables numériques. Ce faisant, il faut garder leur différence ; car, par exemple, $\pi R^2 H \neq \pi H^2 R$. A ces fins il faut disposer les variables dans un ordre déterminé et écrire leurs valeurs comme suit : $(x_1, \ldots, x_n)$ où à la k-ième place on trouve la k-ième variable. Ainsi donc, les ensembles des valeurs envisagées sont des points de l'espace correspondant $\mathbf{R}^n$;

2) faire abstraction du caractère concret des restrictions qui lient les ensembles des variables. Il en reste alors tout seulement un ensemble de $\mathbf{R}^n$;

3) faire abstraction du caractère concret de la dépendance d'une grandeur donnée (envisagée déjà comme variable numérique) par rapport aux éléments de cet ensemble.

Nous aboutissons alors à la notion suivante.

DÉFINITION 1. On appelle *fonction réelle de n variables réelles* une application arbitraire de $\mathbf{R}^n$ dans $\mathbf{R}$. La valeur d'une telle fonction f en un point $(x_1, \ldots, x_n)$ est notée $f(x_1, \ldots, x_n)$, et $x_1, \ldots, x_n$ sont appelés *arguments* de la fonction f.

Si une fonction est donnée par la formule et son ensemble de définition n'est pas explicitement défini, on sous-entend (comme pour la fonction d'une seule variable) que ce dernier représente un ensemble de tous les n-uples d'arguments pour lesquels la formule a un sens.

Par exemple, pour une fonction f donnée par la formule

$$f(x, y, z) = \frac{\ln (x - y)}{\sqrt{x^2 + z^2}},$$

D_f n'étant pas indiqué, on sous-entend que $D_f = \{x, y, z \in \mathbf{R}^3 | \ x > y$ et $x^2 + z^2 \neq 0\}$. Il en est de même pour une fonction qui exprime une grandeur concrète quelconque : on sous-entend que les arguments dont dépend cette dernière ne sont soumis qu'aux restrictions déterminées par leur sens concret.

EXERCICE 1. Montrer que les longueurs des côtés du triangle peuvent être exprimées par des nombres réels quelconques satisfaisant aux inégalités $x + y > z$, $y + z > x$, $x + z > y$. Décrire géométriquement la partie de l'espace des coordonnées, qui est occupée par ces points (x, y, z) (et qui est un ensemble de définition d'une fonction exprimant, par exemple, l'aire du triangle par l'intermédiaire des longueurs de ses côtés).

REMARQUE. On dit souvent que $f(x_1, \ldots, x_n)$ est une fonction de n variables indépendantes. Ceci n'est pas toujours vrai : nous avons vu que les arguments peuvent être liés par une dépendance (ou même par plusieurs dépendances). Cependant, si D_f est un ensemble ouvert, les arguments sont dans une certaine mesure indépendants : si $a = (a_1, \ldots, a_n) \in D_f$, il existe un $\delta > 0$ tel que le cube $Q(a, \delta)$ est contenu dans D_f et, par conséquent, $x_1, \ldots, x_n$ peuvent varier indépendamment l'un de l'autre dans les intervalles correspondants $]a_k - \delta, a_k + \delta[$ sans sortir de D_f.

Les fonctions réelles d'une variable réelle sont représentées dans le plan de coordonnées par un *graphe*, i.e. par un ensemble de points $(x, f(x))$ où x parcourt D_f. De façon analogue, on appelle *graphe d'une fonction f de deux variables* l'ensemble de points $\{(x, y, f(x, y)) | (x, y) \in D_f\}$ de l'espace des coordonnées à trois dimensions. Habituellement, c'est une surface (fig. 12). D'une façon plus générale, le graphe d'une fonction $f(x_1, \ldots, x_n)$ est un ensemble de points $(x_1, \ldots, x_n, f(x_1, \ldots, x_n)) \in \mathbf{R}^{n+1}$ où $(x_1, \ldots \ldots, x_n)$ parcourt D_f.

Nous avons donc défini la fonction de plusieurs variables comme une fonction d'un seul point variable de l'espace $\mathbf{R}^n$. Mais la notion de fonction de plusieurs variables a quelques traits particuliers qui découlent du carac-

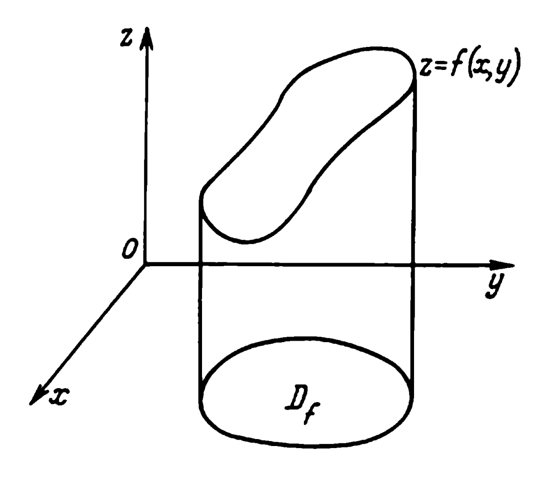


Fig. 12

tère spécifique de l'espace $\mathbf{R}^n$. Ce n'est pas un simple espace normé à n dimensions mais un espace réalisé comme ensemble des n-uples $(x_1, \ldots, x_n)$. Grâce à cela on voit surgir de nouvelles notions.

2. Fonctions de coordonnées

Ce n'est pas seulement le point x de l'espace $\mathbf{R}^n$ qui dépend de ses coordonnées $x_1, \ldots, x_n$, mais, inversement, *les coordonnées du point dépendent de ce point*. A savoir, il existe *n applications* de $\mathbf{R}^n$ dans $\mathbf{R}$ définies par les formules

$$\pi_k(x_1, \ldots, x_n) = x_k \qquad (k = 1, \ldots, n).$$

Les fonctions π_k (« projections sur les axes de coordonnées ») seront appelées *fonctions de coordonnées* sur $\mathbf{R}^n$. Elles admettent, dans un certain sens, des « fonctions réciproques » qui sont les n injections de $\mathbf{R}$ dans $\mathbf{R}^n$ définies par les formules

$$i_k(\lambda) = \lambda e^k \qquad (\lambda \in \mathbf{R}),$$

où e^k est le k-ième vecteur de la base standard dans $\mathbf{R}^n$ (i.e. $\pi_l(e^k) = \delta_l^k$ où δ_l^k est le symbole de Kronecker égal à 1 pour $l = k$ et à 0 pour $l \neq k$). D'où, pour tout $\lambda \in \mathbf{R}$,

$$\pi_k(i_k(\lambda)) = \begin{cases} \lambda & \text{pour } l = k, \\ 0 & \text{pour } l \neq k, \end{cases}$$

i.e.

$$\pi_l \circ i_k = \begin{cases} I_{\mathbf{R}} & \text{pour } l = k, \\ 0 & \text{pour } l \neq k \end{cases}$$

(rappelons que le symbole I_A désigne une application identique de A sur A). D'autre part, pour tout $x \in \mathbf{R}^n$ on a

$$x = (x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k e^k = \sum_{k=1}^{n} i_k(x_k) = \sum_{k=1}^{n} i_k(\pi_k(x)),$$

d'où

$$\sum_{k=1}^{n} i_k \circ \pi_k = I_{\mathbf{R}^n}.$$

3. Fonctions partiellement constantes

A côté des fonctions constantes on rencontre des fonctions *partiellement constantes*, c'est-à-dire les fonctions qui sont constantes par rapport à une partie d'arguments. Telles sont, par exemple, les fonctions de coordonnées. Les fonctions partiellement constantes peuvent apparaître tout naturellement à la suite d'opérations sur les fonctions non partiellement constantes ; par exemple, $(x + y + z) + (x + y - z) = 2(x + y)$, de sorte que dans ce contexte il est naturel d'envisager $2(x + y)$ comme une fonction de trois variables x, y, z, qui est constante par rapport à z.

On peut traiter toute fonction de n variables en tant que fonction de n'importe quel nombre de variables de base en posant par définition $\bar{f}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}) = f(x_1, \ldots, x_n)$, avec $D_{\bar{f}} = D_f \times \mathbf{R}^m$ où $\mathbf{R}^m$ est formé de tous les points $(x_{n+1}, \ldots, x_{n+m})$. Cela devient même inévitable si, par exemple, on veut considérer

$$f(x, y, z) = (x - y)^2 - \ln z$$

comme différence des fonctions $(x - y)^2$ et $\ln z$; il faut alors traiter celles-ci comme fonctions des trois arguments x, y, z (partiellement constantes respectivement par rapport à z ou à x et y). La possibilité d'augmenter ainsi le nombre des arguments n'est pas moins importante lors de l'étude des fonctions « composées ». Or une analyse approfondie de ces dernières s'avère impossible sans qu'on généralise préalablement la notion de fonction numérique de n variables et qu'on introduise une fonction de n variables à valeurs dans un espace de dimension m.

4. Fonction de $\mathbf{R}^n$ dans $\mathbf{R}^m$

DÉFINITION 2. On appelle *fonction de n variables réelles à valeurs dans* $\mathbf{R}^m$ toute application f de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Les *composantes* de cette application sont par définition m fonctions réelles

$$f_k = \pi_k \circ f \quad (k = 1, \ldots, m) \tag{2}$$

où π_k sont les fonctions de coordonnées sur $\mathbf{R}^m$. On définit de la même façon les composantes f_k de l'application f de tout ensemble dans $\mathbf{R}^m$, i.e. $f_k(x) = \pi_k(f(x))$ n'est pas autre chose que la *k-ième coordonnée du vecteur* $f(x)$. Pour cette raison, *toute fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est bien définie par ses composantes.* En vertu des formules (1) et (2), f s'exprime par l'intermédiaire de ses composantes d'après la formule

$$f = \sum_{k=1}^{m} i_k \circ f_k. \tag{3}$$

En effet,

$$f = I_{\mathbf{R}^m} \circ f = \left(\sum_{k=1}^{m} i_k \circ \pi_k\right) \circ f = \sum_{k=1}^{m} i_k \circ \left(\pi_k \circ f\right) = \sum_{k=1}^{m} i_k \circ f_k.$$

Il est évident que

$$D_{f_k} = D_f \quad (k = 1, \ldots, m).$$

Ceci étant, quelles que soient m fonctions réelles f_k de n variables réelles ayant un ensemble de définition D commun, il existe une fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ dont elles sont composantes, i.e. une fonction f définie pour tout $x \in D$ par la formule

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

Nous citerons plus loin quelques exemples importants des fonctions de $\mathbf{R}^n$ dans $\mathbf{R}^m$.

5. Fonction composée

Pour les fonctions d'une variable, la notion de fonction composée coïncide avec celle de composée de fonctions. Ce n'est plus ainsi pour les fonctions de plusieurs variables. Par exemple, soient données les fonctions $g(u, v)$, $\varphi(x, y)$, $\psi(y, z, t)$. En substituant φ à u et ψ à v, on obtient $g(\varphi(x, y), \psi(y, z, t))$. C'est une « fonction composée », mais ce n'est pas une composée de fonctions. Cependant, on peut la transformer en celle-ci. En effet, les arguments u et v de la fonction initiale g représentent des fonctions de différents groupes de variables ; il y a en tout quatre variables : x, y, z, t. Première étape : on transforme φ et ψ en fonctions de ces quatre variables, i.e. on remplace φ et ψ par des fonctions partiellement constantes

$$\bar{\varphi}(x, y, z, t) = \varphi(x, y), \text{ avec } D_{\bar{\varphi}} = D_\varphi \times \mathbf{R}_z \times \mathbf{R}_t,$$
$$\bar{\psi}(x, y, z, t) = \psi(y, z, t), \text{ avec } D_{\bar{\psi}} = \mathbf{R}_x \times D_\psi$$

(où $\mathbf{R}_z$ symbolise bien sûr l'espace $\mathbf{R}$ traité en tant qu'ensemble de variation de l'argument z ; il en est de même pour $\mathbf{R}_t$ et $\mathbf{R}_x$). La substitution de ces fonctions à u et v donne

$$g(\varphi(x, y), \psi(y, z, t)) = g(\bar{\varphi}(x, y, z, t), \bar{\psi}(x, y, z, t)).$$

Deuxième étape : on considère $\bar{\varphi}$ et $\bar{\psi}$ en tant que composantes d'une application de $\mathbf{R}^4$ dans $\mathbf{R}^2$, soit χ, définie par

$$D_\chi = D_{\bar{\varphi}} \cap D_{\bar{\psi}}, \; \chi(x, y, z, t) = (\bar{\varphi}(x, y, z, t), \bar{\psi}(x, y, z, t)).$$

Dans ce cas la fonction « composée » $g(\varphi(x, y), \psi(y, z, t))$ se transforme en composée $g \circ \chi$. Sans introduction de fonctions à valeurs dans $\mathbf{R}^m$ (ici dans $\mathbf{R}^4$), ce serait impossible. Donc, même si l'on n'a en vue que l'étude des fonctions numériques de plusieurs variables scalaires, il devient aussi nécessaire de considérer les fonctions vectorielles.

6. Fonction partielle

Soit $g(x, y, z, t)$ une fonction de quatre variables réelles à valeurs dans $\mathbf{R}^m$. En fixant les valeurs $y = y_0$ et $t = t_0$, on obtient une fonction de deux variables x, z, soit

$$f(x, z) = g(x, y_0, z, t_0) \quad \text{avec } D_f = \{(x, z) \in \mathbf{R}^2 | (x, y_0, z, t_0) \in D_g\},$$

i.e. f est un exemple des *fonctions partielles* qu'on peut obtenir de la fonction g en fixant les valeurs d'un seul, de deux ou de trois de ces arguments. La fonction f est la composée des applications $\mathbf{R}^2 \xrightarrow{h} \mathbf{R}^4 \xrightarrow{g} \mathbf{R}^m$ où $h(x, z) = (x, y_0, z, t_0)$. La fonction de $\mathbf{R}$ dans $\mathbf{R}^m$ qui s'obtient si l'on fixe les valeurs de toutes les coordonnées du point $x = (x_1, \ldots, x_n)$, excepté la k-ième, est appelée *k-ième fonction partielle* de la fonction f de $\mathbf{R}^n$ dans $\mathbf{R}^m$.

§ 2. Limite et continuité d'une fonction de plusieurs variables

Puisque la fonction de n variables réelles à valeurs dans $\mathbf{R}^m$ est définie comme application de $\mathbf{R}^n$ dans $\mathbf{R}^m$ (qui sont des espaces topologiques), les définitions de la limite et de la continuité et leurs propriétés communes à tous les espaces topologiques restent valables dans ce cas particulier. Pour cette raison, nous ne nous arrêterons que sur certains aspects nouveaux qui apparaissent à la suite de l'étude de ces applications en tant que fonctions de plusieurs variables.

1. Quelques exemples de limites

EXEMPLE 1. Soit $f(x, y, z) = xyz/(x^2 + y^2 + z^2)$. Ici $(0, 0, 0) \notin D_f$. Montrons que $f(x, y, z) \to 0$ quand $(x, y, z) \to (0, 0, 0)$. Supposons que $\varrho = \|(x, y, z) - (0, 0, 0)\| = \sqrt{x^2 + y^2 + z^2}$ et que α, β, γ soient les angles que le rayon vecteur du point (x, y, z) forme avec les axes de coordonnées, de sorte que

$$x = \varrho \cos \alpha, \qquad y = \varrho \cos \beta, \qquad z = \varrho \cos \gamma$$

(fig. 13). On a alors $f(x, y, z) = \varrho \cos \alpha \cos \beta \cos \gamma$, d'où $|f(x, y, z)| \leqslant \varrho$, de sorte que $|f(x, y, z)| < \varepsilon$ dès que $\varrho < \varepsilon$.

EXEMPLE 2. Soit $f(x, y) = xy/(x^2 + y^2)$. Ici $(0, 0) \notin D_f$. Supposons que $\varrho = \|(x, y) - (0, 0)\| = \sqrt{x^2 + y^2}$ et que α, β soient des angles formés par le rayon vecteur du point (x, y) avec les axes de coordonnées, de sorte que $x = \varrho \cos \alpha$, $y = \varrho \cos \beta$ $(= \varrho \sin \alpha)$ (fig. 14). On a alors $f(x, y) = \cos \alpha \cos \beta$. Ainsi donc, f est constante sur chaque droite passant par l'origine des coordonnées. Comme α et β ne sont liés que par la condition $\cos^2 \alpha + \cos^2 \beta = 1$, on obtient $|\cos \alpha \cos \beta| \leqslant 1/2$, et le produit

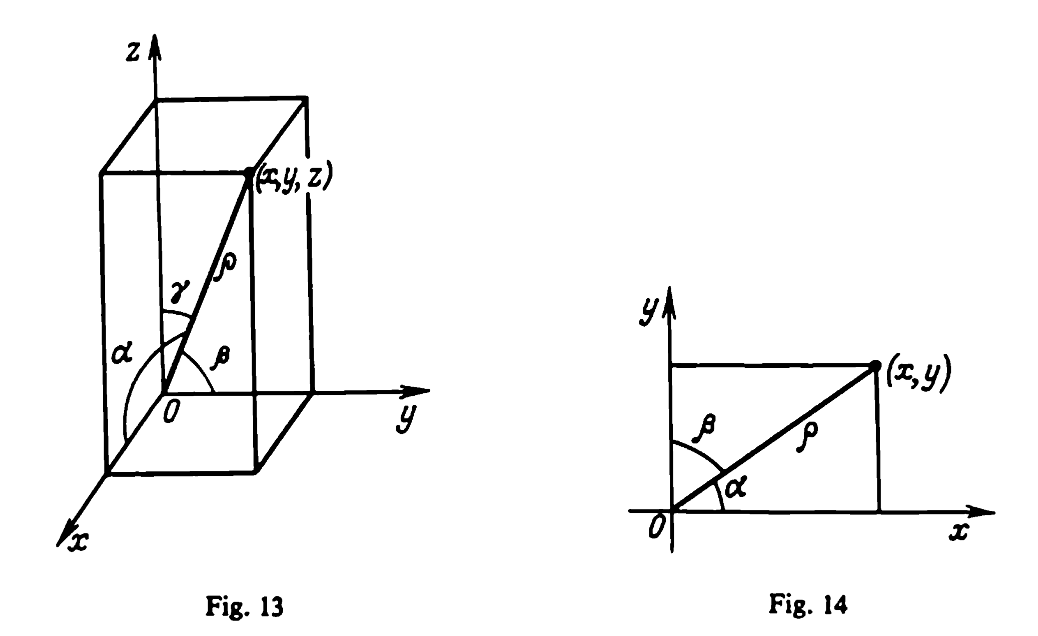


Fig. 13 Fig. 14

cos α cos β peut prendre toute valeur allant de $-1/2$ à $1/2$. Ainsi donc, la fonction f admet une limite lorsque $(x, y) \to (0, 0)$ suivant chaque droite passant par l'origine des coordonnées. Mais ces limites sont distinctes et c'est pourquoi $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ n'existe pas.

EXEMPLE 3. Soit

$$f(x, y) = \begin{cases} 1 & \text{si} \quad 0 < y < x^2, \\ 0 & \text{si} \quad y \leqslant 0 \text{ ou } y \geqslant x^2 \end{cases}$$

(fig. 15). Ici de nouveau la fonction f a une limite quand $(x, y) \to (0, 0)$ suivant chaque droite passant par l'origine des coordonnées. De plus, ces limites coïncident (= 0). Cependant, $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ n'existe pas puisque dans tout voisinage du point (0, 0) il existe un point (x, y) où $f(x, y) = 0$ et un point (x, y) où $f(x, y) = 1$.

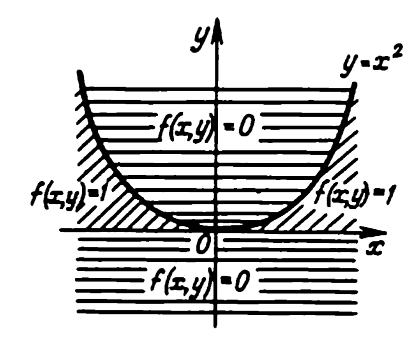


Fig. 15

2. Continuité de certaines applications

LEMME. *Soient X et Y des espaces normés, f une application de X dans Y et $a \in D_f$. S'il existe un nombre C tel que $\|f(x) - f(a)\| \leqslant C\|x - a\|$ pour tout $x \in D_f$, f est continue au point a.*

Démonstration. Si $C = 0$, f est continue, et par suite, partout continue. Soit $C > 0$. Alors pour tout nombre $\varepsilon > 0$ il existe un nombre $\delta > 0$ (par exemple, $\delta = \varepsilon/C$) tel que $\|f(x) - f(a)\| < \varepsilon$ pour tout $x \in D_f$ satisfaisant à l'inégalité $\|x - a\| < \delta$.

Utilisons ce lemme pour établir la continuité de certaines applications importantes.

1° *Continuité de la norme.* Soit X un espace normé. Sa norme vérifie l'inégalité

$$|\|x\| - \|a\|| \leqslant \|x - a\|$$

qui se démontre de la même façon que dans le cas où $X = \mathbf{R}$ et $\|x\| = |x|$. En supposant que $f(x) = \|x\|$ et $Y = \mathbf{R}$, on voit que l'hypothèse du lemme est vérifiée pour $C = 1$. Ainsi donc, $\|x\|$ est une fonction continue de x.

2° *Continuité des injections $i_k : \mathbf{R} \to \mathbf{R}^m$* (voir n° 2 du § 1). Pour tous $\alpha, \lambda \in \mathbf{R}$ on a

$$\|i_k(\lambda) - i_k(\alpha)\| = \|\lambda e^k - \alpha e^k\| = \|(\lambda - \alpha)e^k\| = |\lambda - \alpha|\|e^k\| = |\lambda - \alpha|,$$

de sorte qu'on peut appliquer le lemme pour $X = \mathbf{R}$, $Y = \mathbf{R}^m$, $f = i_k$ et $C = 1$.

3° *Projecteurs.* Toute suite de numéros $i_1, \ldots, i_m$ satisfaisant aux inégalités $1 \leqslant i_1 \leqslant \ldots \leqslant i_m \leqslant n$ engendre une application $\mathrm{pr}_{i_1, \ldots, i_m} : \mathbf{R}^n \to \mathbf{R}^m$ définie d'après la formule

$$\mathrm{pr}_{i_1, \ldots, i_m}(x_1, \ldots, x_n) = (x_{i_1}, \ldots, x_{i_m}).$$

Par exemple, $\mathrm{pr}_{1,3} : \mathbf{R}^4 \to \mathbf{R}^2$ associe à chaque point $(x, y, z, t) \in \mathbf{R}^4$ le point $(x, z) \in \mathbf{R}^2$. Ces applications s'appellent *projecteurs*. En particulier, lorsque $m = 1$, on obtient les fonctions de coordonnées $\mathrm{pr}_k = \pi_k$. Vu que

$$\|\mathrm{pr}_{i_1, \ldots, i_m}(x) - \mathrm{pr}_{i_1, \ldots, i_m}(a)\| =$$
$$= \sqrt{\sum_{k=1}^{m} (x_{i_k} - a_{i_k})^2} \leqslant \sqrt{\sum_{l=1}^{m} (x_l - a_l)^2} = \|x - a\|,$$

les projecteurs sont continus en vertu du lemme (pour $C = 1$). En particulier, *les fonctions de coordonnées sont continues.*

4° *Fonctions rationnelles.* On appelle *fonction rationnelle de plusieurs variables réelles* la fonction qu'on obtient à partir de ces variables et des constantes à l'aide d'un nombre fini d'opérations arithmétiques. *Les fonctions rationnelles de plusieurs variables sont continues.* En effet, toute variable traitée en tant que fonction de l'ensemble de ces variables, i.e. fonc-

tion de coordonnées, est continue (voir propriété 3°) ; les constantes sont aussi continues et les opérations arithmétiques ne violent pas la continuité.

Notamment sont continues toutes les fonctions rationnelles entières, i.e. présentées sous forme de polynômes

$$\sum a_{m_1, \ldots, m_n} x_1^{m_1} \ldots x_n^{m_n}$$

à exposants entiers. Par exemple, le produit scalaire $\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k$ est une fonction continue de $2n$ variables $x_1, \ldots, x_n, y_1, \ldots, y_n$.

5° *Applications dans* $\mathbf{R}^m$ *et leurs composantes.* Soient f une application d'un espace topologique X dans $\mathbf{R}^m$ et $f_1, \ldots, f_m$ ses composantes. Etant donné que $f_k = \pi_k \circ f$ (voir formule (2) du § 1) et les applications $\pi_k : \mathbf{R}^m \to \mathbf{R}$ sont continues (voir propriété 3°), le théorème de la continuité d'une application composée dit que *la continuité de l'application f au point a entraîne celle de toutes ses composantes en ce point.* Inversement : comme $f = \sum_{k=1}^{n} i_k \circ f_k$ (la formule (3) du § 1 reste valable si l'ensemble de départ de f est un ensemble quelconque et non seulement $\mathbf{R}^n$) et les applications $i_k : \mathbf{R} \to \mathbf{R}^m$ sont continues (propriété 2°), *la continuité en a de toutes les composantes de l'application f entraîne celle de f en ce point.* De façon analogue, en vertu du théorème de passage à la limite sous le signe de la fonction continue, $(f(x) \to b = (b_1, \ldots, b_m)$ lorsque $x \to a) \Leftrightarrow (f_k(x) \to b_k$ lorsque $x \to a$ $(k = 1, 2, \ldots, m))$.

6° *Permutation des coordonnées.* On appelle ainsi une application $j : \mathbf{R}^n \to \mathbf{R}$ définie par la formule

$$j(x_1, \ldots, x_n) = (x_{i_1}, \ldots, x_{i_n}),$$

où $i_1, \ldots, i_n$ est une permutation des numéros $1, \ldots, n$. (Exemple : application de $\mathbf{R}^3$ dans lui-même qui fait correspondre à tout point (x, y, z) le point (y, z, x).) Comme $\|j(x) - j(a)\| = \|x - a\|$, *la permutation des coordonnées est continue.*

7° *Fonctions partiellement constantes.* Soit $\bar{f}$ une fonction partiellement constante obtenue de la fonction $f : \mathbf{R}^m \to \mathbf{R}$ par adjonction de p nouvelles variables :

$$D_{\bar{f}} = D_f \times \mathbf{R}^n, \qquad \bar{f}(x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+p}) = f(x_1, \ldots, x_m).$$

Si f est continue au point $a = (a_1, \ldots, a_m)$, $\bar{f}$ est continue en tout point $(a_1, \ldots, a_m, x_{m+1}, \ldots, x_{m+p})$ $((x_{m+1}, \ldots, x_{m+p}) \in \mathbf{R}^p)$. En effet, $\bar{f} = f \circ \mathrm{pr}_{1,\ldots,m}$ où $\mathrm{pr}_{1,\ldots,m}$ est un projecteur de $\mathbf{R}^{m+p}$ sur $\mathbf{R}^m$ défini par la formule

$$\mathrm{pr}_{1,\ldots,m}(x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+p}) = (x_1, \ldots, x_m)$$

et continu d'après la propriété 3°.

REMARQUE. Les nouvelles variables peuvent être disposées partout et non seulement à la fin. Dans ce cas, au lieu de $\text{pr}_{1,\ldots,m}$ il faudrait appliquer le projecteur $\text{pr}_{i_1,\ldots,i_m}$ où $i_1, \ldots, i_m$ sont les numéros des places où se trouveraient les anciennes variables $x_1, \ldots, x_m$.

8° *Fonctions composées.* Il résulte des propriétés 7° et 5° et de la continuité de la composée d'applications continues que *la fonction composée formée de fonctions continues est continue.* Ainsi, soit

$$f(x, y, z, t) = g[\varphi(x, y), \psi(y, z, t)]$$

(voir n° 5 du § 1). *Si φ et ψ sont continues aux points respectifs* (x_0, y_0) et (y_0, z_0, t_0) *et si g est continue au point* (u_0, v_0) *où* $u_0 = \varphi(x_0, y_0)$, $v_0 = \psi(y_0, z_0, t_0)$, *on en conclut que f est continue au point* (x_0, y_0, z_0, t_0). En effet,

$$f(x, y, z, t) = g[\bar{\varphi}(x, y, z, t), \bar{\psi}(x, y, z, t)],$$

où $\bar{\varphi}(x, y, z, t) = \varphi(x, y)$, $\bar{\psi}(x, y, z, t) = \psi(y, z, t)$. Comme φ est continue au point (x_0, y_0), $\bar{\varphi}$ est continue au point (x_0, y_0, z_0, t_0) en vertu de la propriété 7°. De façon analogue, la continuité de ψ au point (y_0, z_0, t_0) entraîne celle de $\bar{\psi}$ au point (x_0, y_0, z_0, t_0). Par conséquent, d'après la propriété 5°, l'application $\chi : \mathbf{R}^4 \to \mathbf{R}^2$ de composantes $\chi_1 = \bar{\varphi}$, $\chi_2 = \bar{\psi}$ est continue en (x_0, y_0, z_0, t_0). D'autre part, étant donné que g est continue au point $(u_0, v_0) = \chi(x_0, y_0, z_0, t_0)$ et que $f = g \circ \chi$, f est continue en (x_0, y_0, z_0, t_0).

9° *Adjonction des valeurs fixées des nouvelles coordonnées.* Soient m, $p \in \mathbf{N}$. Posons

$$h(x_1, \ldots, x_m) = (x_1, \ldots, x_m, x^0_{m+1}, \ldots, x^0_{m+p}),$$

où $x^0_{m+1}, \ldots, x^0_{m+p}$ sont des nombres arbitraires fixés. (Par exemple, associons à chaque point $(x, y) \in \mathbf{R}^2$ le point $(x, y, 1)$ du plan donné par l'équation $z = 1$; fig. 16.) En vertu du lemme, *h est une application continue*

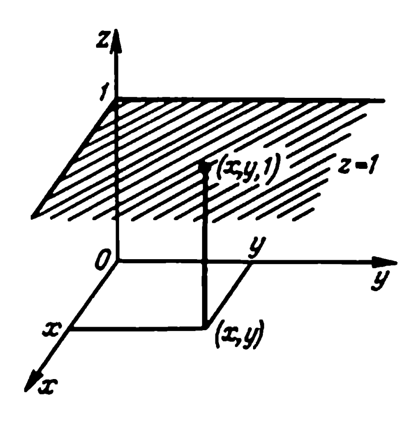


Fig. 16

de $\mathbf{R}^m$ *dans* $\mathbf{R}^{m+p}$ puisque $\|h(x) - h(a)\| = \|x - a\|$. D'une façon plus générale, les valeurs fixées de p nouvelles coordonnées peuvent être mises n'importe où : traitée en tant que composée de l'application h et de la permutation des coordonnées correspondantes, cette application de $\mathbf{R}^m$ dans $\mathbf{R}^{m+p}$ est aussi continue.

10° *Fonctions partielles.* Soit $f(x, z) = g(x, y_0, z, t_0)$. *Si g est continue au point* (x_0, y_0, z_0, t_0), *f est continue au point* (x_0, z_0). En effet, $f = g \circ h$ où h est une application de $\mathbf{R}^2$ dans $\mathbf{R}^4$ définie d'après la formule $h(x, z) = (x, y_0, z, t_0)$; h est continue partout (voir propriété 9°) et g, au point $h(x_0, z_0)$. La même chose est valable pour les autres fonctions partielles.

En particulier, *si $f(x, y)$ est continue au point* (x_0, y_0), *on peut conclure que $f(x, y_0)$ est continue au point x_0 et $f(x_0, y)$ au point y_0*. Notons que la réciproque n'est pas vraie. Soit, par exemple,

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{pour} \quad (x, y) \neq (0, 0), \\ 0 & \text{pour} \quad (x, y) = (0, 0) \end{cases}$$

et $(x_0, y_0) = (0, 0)$. Les fonctions $f(x, 0)$ et $f(0, y)$ sont nulles et, par suite, continues. Cependant, f est discontinue au point $(0, 0)$ (voir exemple 2).

Deuxième partie

APPLICATIONS DIFFÉRENTIABLES D'UN ESPACE FINIDIMENSIONNEL DANS L'AUTRE

CHAPITRE 6

FONCTIONS RÉELLES DIFFÉRENTIABLES

§ 1. Dérivées partielles et dérivées dans une direction

1. Accroissement de l'argument et de la fonction

Vu que $\mathbf{R}^n$ est un espace vectoriel, la notion d'accroissement de l'argument et de la fonction pour les fonctions de plusieurs variables peut être définie de la même façon que pour la fonction d'une variable. Soit f une application de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Fixons un point $x \in D_f$. On appelle *accroissement de l'argument* x de f et on note Δx la différence entre une valeur arbitraire $x' \in D_f$ et x. Ainsi

$$\Delta x = x' - x, \qquad x' = x + \Delta x.$$

La différence entre la valeur de f en x' et la valeur en x est appelée *accroissement de la fonction f au point x*. On la note habituellement $\Delta f(x)$. Ainsi,

$$\Delta f(x) = f(x') - f(x) = f(x + \Delta x) - f(x).$$

Ici Δx et $\Delta f(x)$ sont des vecteurs (resp. de dimension n et m). Si $x = (x_1, \ldots, x_n)$ et $x' = (x_1', \ldots, x_n')$, on a

$$\Delta x = (x_1' - x_1, \ldots, x_n' - x_n) = (\Delta x_1, \ldots, \Delta x_n) = \sum_{k=1}^{n} \Delta x_k e^k$$

(où e^k sont les vecteurs de la base standard dans $\mathbf{R}^n$, voir n° 1.1.1).

On envisage souvent le cas où une seule coordonnée du vecteur x, par exemple la k-ième, est variable. Cela s'exprime comme suit :

$$\Delta x = \Delta x_k e^k = (0, \ldots, 0, \Delta x_k, 0, \ldots, 0)$$

où Δx_k est à la k-ième place et toutes les autres coordonnées sont nulles.

2. Dérivées partielles

DÉFINITION 1. On appelle *dérivée partielle* d'une fonction réelle f de n variables réelles par rapport à la k-ième variable (« par rapport à x_k ») la fonction (notée $\partial_k f$) satisfaisant aux conditions suivantes :

a) son ensemble de définition est l'ensemble des points $x \in D_f$ où le rapport $\dfrac{f(x+\Delta x_k e^k)-f(x)}{\Delta x_k}$ tend vers une limite finie lorsque $\Delta x_k \to 0$;

b) en chacun de ces points la valeur de $\partial_k f$ est égale à cette limite :

$$(\partial_k f)(x) = \lim_{\Delta x_k \to 0} \frac{f(x+\Delta x_k e^k)-f(x)}{\Delta x_k}.$$

Ainsi, par exemple pour une fonction f de deux variables,

$$\partial_1 f(x, y) = \lim_{h \to 0} \frac{f(x+h, y)-f(x, y)}{h},$$

$$\partial_2 f(x, y) = \lim_{k \to 0} \frac{f(x, y+k)-f(x, y)}{k}$$

(h et k sont les notations standard pour les accroissements du premier et du deuxième argument de la fonction de deux variables).

Pour $\partial_k f$ on emploie d'ordinaire les notations classiques f'_{x_k} ou $\dfrac{\partial f}{\partial x_k}$. Notons que $\partial_k f$ est la dérivée de la fonction d'une seule (de la k-ième) variable, à savoir la dérivée de la k-ième fonction partielle au point x_k. En effet, soit

$$\varphi_k(z) = f(x_1, \ldots, x_{k-1}, z, x_{k+1}, \ldots, x_n)$$

la k-ième fonction partielle de z engendrée par la fonction f et le point x. Alors

$$f(x) = \varphi_k(x_k), \qquad f(x+\Delta x_k e^k) = \varphi_k(x_k+\Delta x_k),$$

de sorte que

$$(1) \qquad (\partial_k f)(x_k) = \lim_{\Delta x_k \to 0} \frac{\varphi_k(x_k+\Delta x_k)-\varphi_k(x_k)}{\Delta x_k} = \varphi'_k(x_k).$$

Ainsi donc, le calcul des dérivées partielles de la fonction réelle de plusieurs variables réelles se ramène à celui des dérivées ordinaires de fonctions de **R** dans **R** et, par conséquent, n'exige aucune technique nouvelle.

EXEMPLE 1. Soit $\varrho(x, y, z) = \sqrt{x^2+y^2+z^2}$. Alors

$$(\partial_1 \varrho)(x, y, z) = \frac{x}{\varrho} = \cos\alpha, \qquad (\partial_2 \varrho)(x, y, z) = \frac{y}{\varrho} = \cos\beta,$$

$$(\partial_3 \varrho)(x, y, z) = \frac{z}{\varrho} = \cos\gamma,$$

où α, β, γ sont les angles que le vecteur (x, y, z) forme avec les axes de coordonnées correspondants. L'ensemble de définition de ces dérivées partielles est $D_\varrho \setminus \{(0, 0, 0)\}$.

EXERCICE 1. Vérifier qu'au point $(0, 0, 0)$ les dérivées partielles de la fonction ϱ n'existent pas.

EXERCICE 2. Trouver les dérivées partielles de la fonction

$$\varphi(x, y) = \operatorname{Arctg} \frac{y}{x}.$$

L'existence des dérivées partielles au point n'impose pas des restrictions très rigoureuses à la fonction. Par exemple pour une fonction f de deux variables, l'existence de $(\partial_1 f)(x_0, y_0)$ et de $(\partial_2 f)(x_0, y_0)$ ne caractérise dans une certaine mesure que le comportement de f sur les droites $y = y_0$ et $x = x_0$ dans un voisinage du point (x_0, y_0). Il est à noter que $\partial_1 f$ et $\partial_2 f$ *peuvent aussi exister aux points de discontinuité*. C'est ainsi que pour la fonction

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & ((x, y) \neq (0, 0)), \\ 0 & ((x, y) = (0, 0)) \end{cases}$$

(voir exemple 5.2.2) on a $f(x, 0) \equiv 0$ et $f(0, y) \equiv 0$, de sorte que $\partial_1 f$ et $\partial_2 f$ existent au point $(0, 0)$; cependant f est discontinue en ce point.

3. Dérivées dans une direction

Les dérivées partielles représentent un cas particulier des *dérivées dans une direction*. Tout vecteur unitaire e de $\mathbf{R}^n$ définit une *direction* ; en disant que le point x' se trouve par rapport au point x dans la direction du vecteur unitaire e, on a en vue que $x' = x + \varrho e$ où $\varrho > 0$ (fig. 17). En particulier, les vecteurs unitaires $e^1, \ldots, e^n$ définissent les *directions positives* des axes de coordonnées.

DÉFINITION 2. On appelle *dérivée dans la direction du vecteur unitaire* e d'une fonction réelle f de n variables réelles la fonction (notée $\partial_e f$) vérifiant les conditions suivantes :

a) son ensemble de définition est l'ensemble des points $x \in D_f$ en lesquels le quotient

$$\frac{f(x + te) - f(x)}{t}$$

tend vers une limite finie lorsque $t \to 0$;

b) en chacun de ces points la valeur de $\partial_e f$ est égale à cette limite :

$$(\partial_e f)(x) = \lim_{t \to 0} \frac{f(x + te) - f(x)}{t}.$$

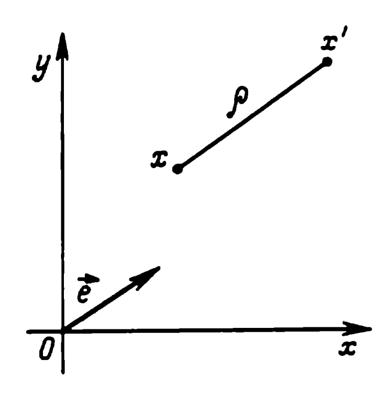


Fig. 17

En particulier, la k-ième dérivée partielle $\partial_k f$ n'est autre que $\partial_{e^k} f$, c'est-à-dire la dérivée dans la direction du k-ième vecteur unitaire e^k (nous avons seulement utilisé pour elle une notation plus brève).

QUESTION. Comment sont liées les dérivées $\partial_{-e} f$ et $\partial_e f$?

Pour exprimer la dérivée dans une direction en coordonnées (et non pas en vecteurs) on peut recourir à une interprétation géométrique des coordonnées du vecteur unitaire en tant que « cosinus directeurs ». Rappelons (voir n° 1.1.6) qu'on appelle angle des vecteurs non nuls x et y dans $\mathbf{R}^n$ l'angle $\varphi \in [0, \pi]$ pour lequel $\cos \varphi = \dfrac{\langle x, y \rangle}{\|x\| \|y\|}$, d'où $\langle x, y \rangle = \|x\| \|y\| \cos \varphi$. Si $x = (x_1, \ldots, x_n)$, on a $x_k = \langle x, e^k \rangle$. Soient $e = (e_1, \ldots, e_n)$ un vecteur unitaire et α_k l'angle des vecteurs e et e^k. On a alors $e_k = \langle e, e^k \rangle = \|e\| \|e^k\| \cos \alpha_k = \cos \alpha_k$, et donc

$$e = (\cos \alpha_1, \ldots, \cos \alpha_n).$$

On dit que $\cos \alpha_1, \ldots, \cos \alpha_n$ sont les *cosinus directeurs* du vecteur unitaire e. Ils satisfont à la relation

$$\sum_{k=1}^{n} \cos^2 \alpha_k = 1 \qquad (= \|e\|)$$

et peuvent être exprimés par n nombres réels quelconques dont la somme des carrés est égale à 1.

On a maintenant $x + te = (x_1 + t \cos \alpha_1, \ldots, x_n + t \cos \alpha_n)$ et, par suite,

$$(2) \quad (\partial_e f)(x) = \lim_{t \to 0} \frac{f(x_1 + t \cos \alpha_1, \ldots, x_n + t \cos \alpha_n) - f(x_1, \ldots, x_n)}{t}.$$

La technique du calcul de cette dérivée sera exposée au § 3.

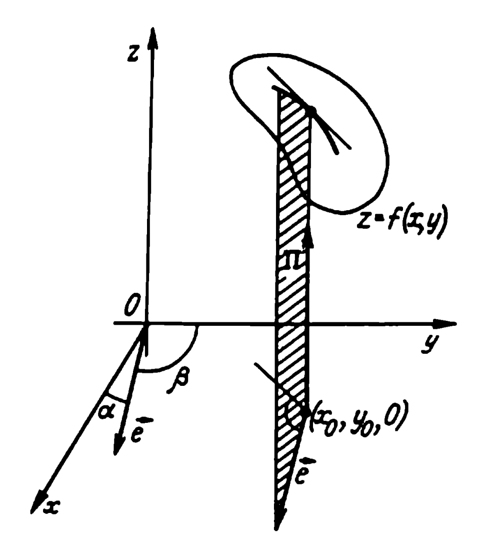


Fig. 18

La formule (2) montre que $(\partial_e f)(x)$ est une dérivée ordinaire d'une fonction d'une variable. En effet, puisque les coordonnées du point x et du vecteur unitaire e sont fixes, posons

$$\varphi(t) = f(x_1 + t \cos \alpha_1, \ldots, x_n + t \cos \alpha_n).$$

Alors

$$(\partial_e f)(x) = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0).$$

Soit, maintenant, f une fonction de deux variables ayant au point (x_0, y_0) une dérivée dans la direction du vecteur unitaire e de cosinus directeurs $\cos \alpha$ et $\cos \beta$ ($= \sin \alpha$). Elucidons le sens géométrique de cette dernière. Portons le vecteur e à partir du point $(x_0, y_0, 0)$ dans le plan $z = 0$ et menons par e le plan Π parallèle à l'axe z. Rapportons Π à un repère orthonormé en prenant e pour le vecteur unitaire de l'axe des abscisses, et la droite passant par $(x_0, y_0, 0)$ parallèlement à l'axe z et ayant le même sens que ce dernier, pour l'axe des ordonnées (fig. 18). Le point $(t, 0)$ du plan Π a, dans le plan (x, y), les coordonnées $x_0 + t \cos \alpha$, $y_0 + t \sin \alpha$, et le plan Π coupe le graphe de la fonction f suivant une courbe qui est, dans Π, le graphe de la fonction

$$\varphi(t) = f(x_0 + t \cos \alpha, x_0 + t \sin \alpha).$$

Puisque φ est dérivable au point $t = 0$, le graphe de φ dans Π a au point correspondant une tangente qui forme avec le vecteur e un angle dont la tangente est égale à $\varphi'(0)$. Or $\varphi'(0) = (\partial_e f)(x_0, y_0)$. Ainsi donc, $(\partial_e f)(x_0, y_0)$ est égal à la tangente de l'angle que forment le plan (x, y) et la tangente en $(x_0, y_0, f(x_0, y_0))$ à la section du graphe de f par le plan Π.

La fonction d'une variable réelle qui a une dérivée en un point donné est continue en ce point. D'autre part, on a déjà vu que pour une fonction de plusieurs variables l'existence des dérivées partielles en un point donné n'entraîne pas, à elle seule, la continuité. Plus encore, l'existence des dérivées dans toutes les directions ne garantie pas la continuité. Ainsi, on a vu que la fonction

$$f(x, y) = \begin{cases} 1 & (0 < y < x^2), \\ 0 & (y \leqslant 0 \text{ ou } y \geqslant x^2) \end{cases}$$

envisagée dans l'exemple 5.2.3 est discontinue au point (0, 0). Cependant, sur chaque droite passant par le point (0, 0) il existe un voisinage de ce point dans lequel f est constante, d'où il résulte qu'au point (0, 0) f a une dérivée (égale à zéro) dans toute direction.

Ainsi donc, l'existence, pour une fonction de plusieurs variables, des dérivées en un point dans toutes les directions est plus faible que l'existence de la dérivée en un point pour une fonction d'une seule variable. Or l'existence de la dérivée pour une fonction d'une seule variable est équivalente à la dérivabilité, i.e. à la « presque linéarité » de l'accroissement de la fonction. La notion de fonction linéaire se généralise à tous les espaces vectoriels, en particulier, à $\mathbf{R}^n$. Cela permet aussi de généraliser, dans $\mathbf{R}^n$, la notion de fonction dérivable.

§ 2. Fonctions linéaires

1. Notion de fonction linéaire

Etant donné un espace vectoriel X sur un corps commutatif K, on appelle *fonction linéaire* sur X toute application linéaire de X dans K, i.e. toute fonction $l : X \to K$ satisfaisant aux conditions

$$l(x + y) = l(x) + l(y) \qquad \text{pour tout } (x, y) \in X \times X,$$
$$l(\lambda x) = \lambda l(x) \qquad \text{pour tout } (x, \lambda) \in X \times K.$$

Il découle de ces conditions que

$$l\left(\sum_{i=1}^{m} \lambda_i x^{(i)}\right) = \sum_{i=1}^{m} \lambda_i l(x^{(i)}).$$

Du cours d'algèbre on sait que quels que soient les espaces vectoriels X et Y sur un même corps commutatif K, les applications linéaires de X dans Y forment un espace vectoriel sur K pour les opérations ordinaires d'addition et de multiplication par des « scalaires » (éléments de K). En particulier, pour un espace vectoriel X sur K, la somme de deux fonctions linéaires sur X et le produit de la fonction linéaire par un scalaire sont des fonctions linéaires, et l'ensemble de toutes les fonctions linéaires sur

X muni de ces opérations est un espace vectoriel sur K ; nous le noterons $\mathscr{L}(X, K)$.

2. Fonctions linéaires sur $\mathbf{R}^n$

Envisageons maintenant d'une façon plus détaillée les fonctions linéaires sur $\mathbf{R}^n$. Au lieu de $l(x)$, nous écrirons, de règle, l_x. Nous avons déjà rencontré certaines fonctions linéaires importantes sur $\mathbf{R}^n$, notamment les fonctions de coordonnées π_k. Rappelons (voir n° 5.1.2) que ces fonctions sont définies par les formules

$$\pi_k(x_1, \ldots, x_n) = x_k \qquad (k = 1, \ldots, n),$$

donc, en particulier,

$$\pi_k(e^i) = \delta_k^i \qquad (i, k = 1, \ldots, n)$$

où δ_k^i sont les symboles de Kronecker. Puisque l'addition des vecteurs et la multiplication du vecteur par un nombre s'effectuent dans $\mathbf{R}^n$ « coordonnée à coordonnée », on a pour tous vecteurs $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ et tout nombre $\lambda \in \mathbf{R}$

$$\pi_k(x + y) = x_k + y_k = \pi_k x + \pi_k y,$$
$$\pi_k(\lambda x) = \lambda x_k = \lambda \pi_k x \ (k = 1, 2, \ldots, n).$$

Cela signifie que *les fonctions de coordonnées π_k sont linéaires.*

THÉORÈME 1. *Les fonctions de coordonnées forment une base dans $\mathscr{L}(\mathbf{R}^n, \mathbf{R})$.*

Démonstration. 1) Si $\lambda_1, \ldots, \lambda_n$ sont des nombres tels que $\sum_{k=1}^{n} \lambda_k \pi_k = 0$, i.e. $\left(\sum_{k=1}^{n} \lambda_k \pi_k\right) x = \sum_{k=1}^{n} \lambda_k \pi_k x = 0$ pour tout $x \in \mathbf{R}^n$, on obtient $\lambda_i = \sum_{k=1}^{n} \lambda_k \delta_k^i = \sum_{k=1}^{n} \lambda_k \pi_k e^i = 0$ pour tout i. Donc, les fonctions de coordonnées π_k sont linéairement indépendantes.

2) Soit $l \in \mathscr{L}(\mathbf{R}^n, \mathbf{R})$. Posons

(1) $$le^k = l_k \ (k = 1, \ldots, n).$$

Pour tout $x = (x_1, \ldots, x_n)$ on a

$$lx = l\left(\sum_{k=1}^{n} x_k e^k\right) = \sum_{k=1}^{n} x_k le^k = \sum_{k=1}^{n} l_k x_k.$$

Or $x_k = \pi_k x$. Par conséquent,

$$lx = \sum_{k=1}^{n} l_k \pi_k x = \left(\sum_{k=1}^{n} l_k \pi_k\right) x$$

pour tout $x \in \mathbf{R}^n$, i.e.

(2) $$l = \sum_{k=1}^{n} l_k \pi_k,$$

de sorte que toute fonction linéaire l sur $\mathbf{R}^n$ est une combinaison linéaire des fonctions de coordonnées π_k.

Ainsi donc, $\mathscr{L}(\mathbf{R}^n, \mathbf{R})$ *est un espace vectoriel réel de dimension n*. Ceci étant, $l \mapsto e$, *où e est un vecteur de* $\mathbf{R}^n$ *de coordonnées* $e_1, \ldots, e_n$, *définit un isomorphisme de* $\mathscr{L}(\mathbf{R}^n, \mathbf{R})$ *sur* $\mathbf{R}^n$. Nous avons par là même obtenu la formule

(3) $$lx = \sum_{k=1}^{n} l_k x_k$$

pour tout $x \in \mathbf{R}^n$ et tout $l \in \mathscr{L}(\mathbf{R}^n, \mathbf{R})$. Notons ses deux corollaires :

1) *toute fonction linéaire sur* $\mathbf{R}^n$ *est continue* (voir n° 5.2.2, 4°) ;

2) $l_x = \langle \vec{e}, x \rangle$.

Dans le cas unidimensionnel (i.e. pour $n = 1$) $\vec{e}$ et x sont des nombres, et $\langle \vec{e}, x \rangle = \vec{e} \cdot x$, de sorte que $l_x = \vec{e} \cdot x$. Ainsi donc, *la fonction linéaire l sur* **R** *est un opérateur de multiplication par le nombre* $\vec{e}$. Dans ce cas il serait souvent rationnel de ne pas faire distinction entre e et $\vec{e}$, i.e. *envisager chaque nombre comme opérateur de multiplication par ce nombre.*

§ 3. Fonctions réelles différentiables de n variables réelles

1. Notion de fonction différentiable de $\mathbf{R}^n$ dans $\mathbf{R}$

DÉFINITION 1. On dit que la fonction f de $\mathbf{R}^n$ dans $\mathbf{R}$ est *différentiable en un point* x si D_f est un voisinage du point x et l'accroissement de la fonction f en ce point peut être présenté sous la forme

$$f(x + \Delta x) - f(x) = l\Delta x + \alpha(x, \Delta x)\|\Delta x\|,$$

où l est une fonction linéaire sur $\mathbf{R}^n$ et $\alpha(x, \Delta x) \to 0$ lorsque $\Delta x \to 0$; l est appelée *dérivée* de f en x et notée $f'(x)$. Ainsi donc,

(1) $$f(x + \Delta x) - f(x) = f'(x)\Delta x + \alpha(x, \Delta x)\|\Delta x\|.$$

La fonction f est dite *différentiable* si elle l'est en tout point $x \in D_f$.

AVERTISSEMENT : $f'(x)$ est une fonction linéaire du vecteur variable h de l'espace $\mathbf{R}^n$, et non pas (en général) de x !

REMARQUE 1. Le sens de la définition 1 est le même que dans le cas unidimensionnel : la différentiabilité de la fonction au point signifie que l'accroissement de la fonction en ce point est « presque linéaire », i.e. est

une fonction linéaire de Δx à un infiniment petit près par rapport à l'accroissement Δx de l'argument.

REMARQUE 2. La définition 1 impose à x une restriction plus forte que celle de la définition de la dérivabilité en x de la fonction f d'une seule variable : on exige que x soit un point non isolé de D_f s'il s'agit d'une fonction d'une seule variable, et qu'il soit contenu à l'intérieur de D_f s'il s'agit d'une fonction de plusieurs variables (d'où il résulte en particulier que D_f est un ensemble ouvert si f est différentiable en tout $x \in D_f$). La commodité de cette restriction devient plus claire par la suite. Elle est dictée par le caractère plus complexe de la topologie du plan en comparaison avec la topologie de la droite.

Voici deux exemples très simples de fonctions différentiables.

EXEMPLE 1. Soit f une fonction constante sur $\mathbf{R}^n$ ou sur un ouvert non vide de $\mathbf{R}^n$. Comme

$$f(x + \Delta x) - f(x) = 0 = 0\Delta x + 0\|x\|,$$

f est différentiable et $f'(x) \equiv 0$.

EXEMPLE 2. Soit l une fonction linéaire sur $\mathbf{R}^n$. Comme

$$l(x + \Delta x) - l(x) = l\Delta x + 0\|\Delta x\|,$$

l est différentiable et

$$l'(x) = l \text{ pour tout } x \in \mathbf{R}^n,$$

ce qui est conforme à la formule $\dfrac{\mathrm{d}(kx)}{\mathrm{d}x} = k$ de la dérivation de la fonction linéaire kx sur $\mathbf{R}$.

THÉORÈME 1. *Si les fonctions f et g de $\mathbf{R}^n$ dans $\mathbf{R}$ sont différentiables au point x, leur somme $f + g$ est différentiable en ce point, et*

$$(f + g)'(x) = f'(x) + g'(x).$$

Démonstration. Soit $h = f + g$. Comme D_f et D_g sont les voisinages du point x, $D_{f+g} = D_f \cap D_g$ est un voisinage de ce point. On a par ailleurs

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + \alpha(x, \Delta x)\cdot\|\Delta x\|, \tag{2}$$

$$g(x + \Delta x) - g(x) = g'(x)\Delta x + \beta(x, \Delta x)\cdot\|\Delta x\|, \tag{3}$$

où $f'(x)$ et $g'(x)$ sont des fonctions linéaires, et $\alpha(x, \Delta x)$ et $\beta(x, \Delta x)$ tendent vers zéro lorsque $\Delta x \to 0$. En additionnant (2) et (3), on obtient

$$h(x + \Delta x) - h(x) = [f'(x) + g'(x)]\Delta x + \gamma(x, \Delta x)\cdot\|\Delta x\|,$$

où $\gamma(x, \Delta x) = \alpha(x, \Delta x) + \beta(x, \Delta x) \to 0$ quand $\Delta x \to 0$. Comme $f'(x) +$ $+ g'(x)$ est une fonction linéaire, on conclut que h est différentiable en x et que $h'(x) = f'(x) + g'(x)$.

De la définition 1 découle facilement le théorème suivant.

THÉORÈME 2. *Si la fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *est différentiable au point* x, *la fonction* λf *est différentiable en ce point pour tout* $\lambda \in \mathbf{R}$, *et* $(\lambda f)'(x) = \lambda f'(x)$.

EXERCICE 1. Montrer que si les fonctions f et g de $\mathbf{R}^n$ dans $\mathbf{R}$ sont différentiables au point x, leur produit est différentiable en ce point. Par quelle formule s'exprime sa dérivée ?

THÉORÈME 3. *La fonction* f *de* $\mathbf{R}^n$ *dans* $\mathbf{R}$ *différentiable au point* x *est continue en ce point.*

Démonstration. Etant une fonction linéaire sur $\mathbf{R}^n$, $f'(x)$ est continue et $f'(x)(0) = 0$. Par conséquent, $f'(x)\Delta x \to 0$ lorsque $\Delta x \to 0$. D'autre part, d'après la définition, $\alpha(x, \Delta x) \to 0$ quand $\Delta x \to 0$ et donc *a fortiori* $\alpha(x, \Delta x)\cdot\|\Delta x\| \to 0$ lorsque $\Delta x \to 0$. Il résulte alors de la formule (1) que $f(x + \Delta x) - f(x) \to 0$ quand $\Delta x \to 0$, i.e. f est continue au point x.

THÉORÈME 4. *La fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *est différentiable au point* $x \in]D_f[$ *si et seulement si l'accroissement de* f *en ce point se présente sous la forme*

$$f(x + \Delta x) - f(x) = l\Delta x + \sum_{k=1}^{n} \alpha_k(x, \Delta x)\Delta x_k \tag{1'}$$

où l *est une fonction linéaire sur* $\mathbf{R}^n$ *et* $\alpha_k(x, \Delta x) \to 0$ *lorsque* $\Delta x \to 0$ *pour tout entier* $k \in [1, n]$.

Démonstration. Soit f une fonction différentiable en x, de sorte qu'on a la formule (1) où $l = f'(x)$ est une fonction linéaire sur $\mathbf{R}^n$ et $\alpha(x, \Delta x) \to 0$ quand $\Delta x \to 0$. Or

$$\alpha(x, \Delta x)\|\Delta x\| = \begin{cases} \displaystyle\sum_{k=1}^{n} \frac{\alpha(x, \Delta x)}{\|\Delta x\|}\Delta x_k^2 & \text{pour} \quad \Delta x \neq 0, \\ 0 & \text{pour} \quad \Delta x = 0. \end{cases}$$

Par conséquent, le second membre de la formule (1) est de la forme (1′) où

$$\alpha_k(x, \Delta x) = \begin{cases} \dfrac{\alpha(x, \Delta x)}{\|\Delta x\|}\Delta x_k & \text{pour} \quad \Delta x \neq 0, \\ 0 & \text{pour} \quad \Delta x = 0. \end{cases}$$

Ceci étant, $\alpha_k(x, \Delta x) \to 0$ quand $\Delta x \to 0$ vu que $|\alpha_k(x, \Delta x)| \leqslant |\alpha(x, \Delta x)|$.

Inversement : soit à démontrer que la fonction f est différentiable en x. Posons

$$\alpha(x, \Delta x) = \begin{cases} \dfrac{1}{\|\Delta x\|}\displaystyle\sum_{k=1}^{n} \alpha_k(x, \Delta x)\Delta x_k & \text{pour} \quad \Delta x \neq 0, \\ 0 & \text{pour} \quad \Delta x = 0. \end{cases}$$

On a alors la formule (1 ′), et $\alpha(x, \Delta x) \to 0$ lorsque $\Delta x \to 0$ car, en vertu de l'inégalité de Cauchy,

$$|\alpha(x, \Delta x)| \leqslant \frac{1}{\|\Delta x\|} \sqrt{\sum_{k=1}^{n} \alpha_k^2(x, \Delta x)} \sqrt{\sum_{k=1}^{n} \Delta x_k^2} = \sqrt{\sum_{k=1}^{n} \alpha_k^2(x, \Delta x)}.$$

2. Différentiabilité et dérivée dans une direction

THÉORÈME 5. *La fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *différentiable au point* x *admet en ce point des dérivées dans toutes les directions, et pour chaque vecteur unitaire* e *de* $\mathbf{R}^n$ *on a*

$$(\partial_e f)(x) = f'(x)e.$$

Démonstration. Comme $x \in]D_f[$, il existe un $\eta > 0$ tel que $\|h\| < \eta \Rightarrow x + h \in D_f$. Tout vecteur h peut être pris pour Δx. En particulier, puisque $\|te\| = |t|$, on peut poser $\Delta x = te$ dans (1) pour tout t tel que $|t| < \eta$. Il s'ensuit que

$$f(x + te) - f(x) = f'(x)(te) + \alpha(x, te)\|te\| = tf'(x)e + \alpha(x, te)|t|$$

d'où

$$\frac{f(x + te) - f(x)}{t} = f'(x)e + \alpha(x, te)\frac{|t|}{t}.$$

Puisque $\alpha(x, te) \to 0$ lorsque $t \to 0$, et $|t|/t$ est borné, on conclut que

$$\frac{f(x + te) - f(x)}{t} \to f'(x)e \quad \text{quand} \quad t \to 0,$$

i.e. la dérivée $(\partial_e f)(x)$ existe et est égale à $f'(x)e$.

COROLLAIRE. *La fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *différentiable au point* x *admet en ce point des dérivées partielles par rapport à tous les arguments, et*

$$(4) \qquad (\partial_k f)(x) = \frac{\partial f}{\partial x_k}(x) = f'(x)e^k \qquad (k = 1, \ldots, n).$$

En vertu des formules (1) et (2) du § 2, $f'(x)e^k$ est le coefficient de π_k dans le développement de la fonction linéaire $f'(x)$ suivant les fonctions de coordonnées. D'après la formule (3) du § 2 on a alors

$$(5) \qquad f'(x)h = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x)h_k \quad \text{pour tout } h \in \mathbf{R}^n.$$

C'est pourquoi l'accroissement de la fonction f différentiable au point x peut être écrit comme suit :

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x)\Delta x_k + \alpha(x, \Delta x)\sqrt{\sum_{k=1}^{n} \Delta x_k^2}$$

où $\alpha(x, \Delta x) \to 0$ quand $\Delta x \to 0$, ou bien, d'après le théorème 4,

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x)\Delta x_k + \sum_{k=1}^{n} \alpha_k(x, \Delta x)\Delta x_k$$

où $\alpha_k(x, \Delta x) \to 0$ lorsque $\Delta x \to 0$ pour tout entier $k \in [1, n]$, ou encore sous la forme

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{n} g_k(x, \Delta x)\Delta x_k$$

où $g_k(x, \Delta x) = \dfrac{\partial f}{\partial x_k}(x) + \alpha_k(x, \Delta x)$, de sorte que

$$g_k(x, \Delta x) \to \frac{\partial f}{\partial x_k}(x) \quad \text{quand} \quad \Delta x \to 0.$$

L'expression (5) est appelée *différentielle* de la fonction f au point x. Elle ne se distingue de la dérivée $f'(x)$ que par la forme car $f'(x)h$ n'est qu'une autre notation de la fonction linéaire $f'(x)$ pour h variable (tout comme $\varphi(z)$ est une autre notation de la fonction φ).

Soit, en particulier, h le vecteur unitaire e de cosinus directeurs $\cos \alpha_1$, $\cos \alpha_2, \ldots, \cos \alpha_n$. On a alors

$$(6) \qquad (\partial_e f)(x)\ (= f'(x)e) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x) \cos \alpha_k.$$

3. Gradient

Dans le n° 2 du § 2 on a associé à toute fonction linéaire l sur $\mathbf{R}^n$ un vecteur $\vec{l}$ dans $\mathbf{R}^n$ de coordonnées $l_k = le^k$ $(k = 1, \ldots, n)$. Il résulte en particulier de la formule (4) que

$$\overrightarrow{f'(x)} = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right).$$

On dit que ce vecteur est le *gradient* de la fonction f au point x et on le note grad $f(x)$. La formule (6) peut alors être mise sous la forme suivante :

$$(\partial_e f)(x) = \langle \operatorname{grad} f(x), e \rangle.$$

Ceci montre que si $\operatorname{grad} f(x) \neq 0$, $\partial_e f$ *présente la plus grande valeur pour* $e = \dfrac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}$, i.e. *lorsque* e *est le vecteur directeur unitaire du gradient.* En effet, d'après l'inégalité de Cauchy-Bouniakovski,

$$|(\partial_e f)(x)| = |\langle \operatorname{grad} f(x), e\rangle| \leqslant \|\operatorname{grad} f(x)\|,$$

d'autre part, $\left\langle \operatorname{grad} f(x), \dfrac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}\right\rangle = \|\operatorname{grad} f(x)\|$, de sorte que $\|\operatorname{grad} f(x)\|$ est la plus grande valeur de la fonction $e(\partial_e f)(x)$, et cette valeur se présente quand e est le vecteur unitaire du vecteur $\operatorname{grad} f(x)$. En interprétant la dérivée $(\partial_e f)(x)$ de la fonction f dans la direction du vecteur unitaire e au point x comme vitesse de variation de f dans cette direction, on peut dire que *le gradient de la fonction en un point est un vecteur qui définit la direction et la vitesse d'accroissement maximal de la fonction en ce point.* Notons encore que

$$\|\operatorname{grad} f(x)\| = \sqrt{\sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(x)\right)^2}.$$

4. Conditions suffisantes de différentiabilité

THÉORÈME 6. *Supposons que la fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *soit définie dans un voisinage* U *du point* x *et qu'elle y possède des dérivées partielles* $\partial_k f$ $(k = 1, \ldots, n)$. *Si elles sont continues au point* x, *la fonction* f *est différentiable en ce point.*

Démonstration. Il existe un $\delta > 0$ tel que $U(x, \delta) \subset U$. Soit $\|\Delta x\| < \delta$. Posons

$$\Delta^0 = 0,\ \Delta^k = \sum_{i=1}^{k} \Delta x_i e^i = (\Delta x_1, \ldots, \Delta x_k, 0, \ldots 0) \quad (k = 1, 2, \ldots, n)$$

(voir fig. 16, où les vecteurs Δ_k sont représentés pour le cas $n = 3$). Comme $\|\Delta^k\| = \sqrt{\sum_{i=1}^{k} \Delta x_i^2} \leqslant \|\Delta x\|$, on a $x + \Delta^k \in U(x, \delta)$. Compte tenu de ce que $\Delta^n = \Delta x$, on obtient

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{n} [f(x + \Delta^k) - f(x + \Delta^{k-1})]$$

et, par conséquent,

$$(7)\quad f(x + \Delta x) - f(x) - \sum_{k=1}^{n} (\partial_k f)(x)\Delta x_k =$$
$$= \sum_{k=1}^{n} [f(x + \Delta^k) - f(x + \Delta^{k-1}) - (\partial_k f)(x)\Delta x_k].$$

Considérons maintenant le segment $[x + \Delta^{k-1}, x + \Delta^k]$. Il est composé des points $x + \Delta^{k-1} + te^k$, où t varie de 0 à Δx_k, et est contenu dans $U(x, \delta)$. Ses extrémités appartiennent à $U(x, \delta)$ qui est convexe. Ainsi donc, f est définie sur ce segment et y possède la dérivée partielle $\partial_k f$. Ayant posé

$$\varphi_k(t) = f(x + \Delta^{k-1} + te^k) = \\ = f(x_1 + \Delta x_1, \ldots, x_{k-1} + \Delta x_{k-1}, t, x_{k+1}, \ldots, x_n),$$

on obtient que φ_k est dérivable et $\varphi_k'(t) = (\partial_k f)(x + \Delta^{k-1} + te^k)$.

Par conséquent, il existe un $\theta_k \in]0, 1[$ tel que

$$(8) \quad f(x + \Delta^k) - f(x + \Delta^{k-1}) = \varphi_k(x_k + \Delta x_k) - \varphi_k(x_k) = \\ = \varphi_k'(x_k + \theta_k \Delta x_k)\Delta x_k = (\partial_k f)(z^k)\Delta x_k$$

où $z^k = x + \Delta^{k-1} + \theta_k \Delta x_k e^k \in U(x, \delta)$. Comme

$$\|z^k - x\| = \|\Delta^{k-1} + \theta_k \Delta x_k e^k\| = \sqrt{\sum_{i=1}^{k-1} \Delta x_i^2 + \theta_k^2 \Delta x_k^2} \leqslant \|\Delta x\|$$

et les fonctions $\partial_k f$ sont continues au point x, on a $(\partial_k f)(z^k) \to (\partial_k f)(x)$ lorsque $\Delta x \to 0$.

Compte tenu de (7) et (8), il vient

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{n} (\partial_k f)(x)\Delta x_k + \sum_{k=1}^{n} \alpha_k(x, \Delta x)\Delta x_k$$

où $\alpha_k(x, \Delta x) \to 0$ lorsque $\Delta x \to 0$ pour tout entier $k \in [1, n]$. Par conséquent, f est différentiable au point x suivant le théorème 4.

COROLLAIRE. *Toute fonction de* $\mathbf{R}^n$ *dans* $\mathbf{R}$ *définie sur un ouvert et ayant partout sur cet ensemble des dérivées partielles continues par rapport à tous les arguments est différentiable.*

5. Interprétation géométrique de la différentielle d'une fonction de deux variables

DÉFINITION 2. Etant donné une fonction $f(x, y)$ continue en un point $P_0(x_0, y_0)$ intérieur à D_f, on appelle *plan tangent* au graphe de f au point $M_0(x_0, y_0, f(x_0, y_0))$ le plan qui passe par M_0 de telle façon que la distance MN d'un point $M(x, y, f(x, y))$ du graphe à ce plan est infiniment petite par rapport à M_0M lorsque $P(x, y) \to P_0$ (fig. 19).

THÉORÈME 7. *Si la fonction* $f(x, y)$ *est différentiable au point* $P_0(x_0, y_0)$, *le graphe* $z = f(x, y)$ *de cette fonction possède au point* $M_0(x_0, y_0, z_0)$, *où* $z_0 = f(x_0, y_0)$, *un plan tangent. Ce plan est défini par l'équation*

$$(9) \qquad Z - z_0 = (\partial_1 f)(x_0, y_0)(x - x_0) + (\partial_2 f)(x_0, y_0)(y - y_0),$$

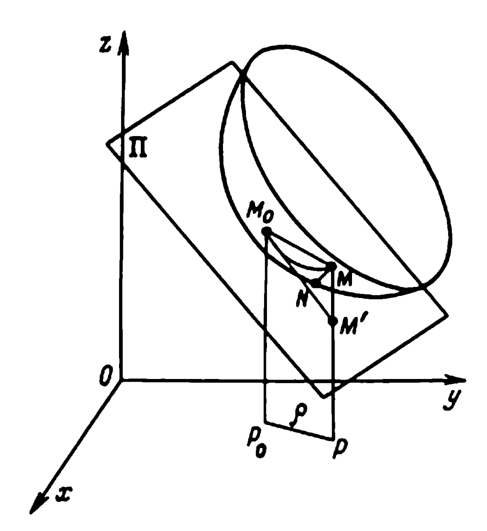


Fig. 19

*de sorte que la valeur de la différentielle de f en P_0 pour les accroissements $x - x_0$, $y - y_0$ des arguments est égale à l'accroissement $Z - z_0$ de la cote d'un point du plan tangent**).

Démonstration. Soient $P(x, y) \in D_f$ et

$$\varrho = P_0P = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

(cf. fig. 19). La fonction f étant différentiable au point P_0, on a

$$(10) \quad z - z_0 = (\partial_1 f)(x_0, y_0)(x - x_0) + (\partial_2 f)(x_0, y_0)(y - y_0) + \alpha(x, y)\cdot\varrho$$

où $\alpha(x, y) \to 0$ lorsque $P \to P_0$, i.e. quand $\varrho \to 0$. D'autre part, l'équation (9) regardée par rapport au point « courant » (x, y, Z) est une équation du plan (noté Π) passant par le point M_0. En retranchant (9) de (10), on obtient $z - Z = \alpha\varrho$. Soit N la base de la perpendiculaire abaissée du point $M(x, y, f(x, y))$ du graphe sur le plan Π et soit M' un point de ce plan ayant les mêmes abscisse et ordonnée que celles de M (et, par suite, la cote Z). Comme $NM \leqslant |M'M|$ et $M_0M \geqslant P_0P$, il vient

$$0 \leqslant \frac{NM}{M_0M} \leqslant \frac{|M'M|}{P_0P} = \frac{|z - Z|}{\varrho} = |\alpha|.$$

Par conséquent, $\dfrac{NM}{M_0M} \to 0$ lorsque $P \to P_0$, i.e. Π est le plan tangent au graphe de la fonction f au point M_0. Ceci étant, le second membre de

*) Les coordonnées courantes d'un point du plan tangent sont désignées ici par x, y, Z à la différence des coordonnées courantes x, y, z d'un point de la surface $z = f(x, y)$.

l'équation (9) de ce plan représente la valeur de la différentielle de la fonction f au point P_0 associée aux valeurs $x - x_0$, $y - y_0$ des accroissements des arguments, et le premier membre est la différence des cotes des points M' et M_0, i.e. l'accroissement de la cote d'un point du plan tangent lors du passage de P_0 à P.

CHAPITRE 7

APPLICATIONS DIFFÉRENTIABLES

§ 1. Applications linéaires

La dérivation des fonctions composées exige l'utilisation des applications différentiables de $\mathbf{R}^n$ dans $\mathbf{R}^m$. La définition de ces dernières repose sur la notion d'application linéaire.

1. Notion d'application linéaire

Soient X et Y des espaces vectoriels sur un même corps commutatif K. On appelle *application linéaire* de X dans Y toute application $f : X \to Y$ satisfaisant aux conditions suivantes :

$$f(x + y) = f(x) + f(y) \qquad \text{pour tout } (x, y) \in X \times X,$$
$$f(\lambda x) = \lambda f(x) \qquad \text{pour tout } (x, \lambda) \in X \times K.$$

En particulier, la fonction linéaire sur X (voir n° 6.2.1) est une application linéaire de X dans K (où K est traité comme espace vectoriel sur lui-même). Dans les considérations générales, les applications linéaires sont habituellement désignées par les majuscules A, B, ... et au lieu de $A(x)$, on écrit Ax.

Du fait que les opérations d'addition et de multiplication par un scalaire sont accomplies coordonnée à coordonnée dans $\mathbf{R}^n$ il résulte facilement que *les projecteurs* $\mathrm{pr}_{i_1 \ldots i_m} : \mathbf{R}^n \to \mathbf{R}^m$ (voir n° 5.2.2, 3°) *et les injections* $i_k : \mathbf{R} \to \mathbf{R}^m$ (voir n° 5.1.2) *sont des applications linéaires.*

Rappelons certains faits concernant les applications linéaires, étudiés dans le cours d'algèbre. Comme nous l'avons déjà dit au n° 6.2.1, les applications linéaires de X dans Y forment un espace vectoriel sur K par rapport aux opérations ordinaires d'addition et de multiplication par un scalaire ; on le note $\mathscr{L}(X, Y)$. Soient X, Y et Z des espaces vectoriels sur K. La composée $B \circ A$ des applications linéaires $A : X \to Y$ et $B : Y \to Z$ est une application linéaire de X dans Z ; au lieu de $B \circ A$ on écrit habituellement BA. La composition des applications linéaires (comme de toutes applications en général) est associative. En outre, BA, traitée en tant que fonction de A et B, est linéaire :

$$B(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 BA_1 + \lambda_2 BA_2,$$
$$(\mu_1 B_1 + \mu_2 B_2)A = \mu_1 B_1 A + \mu_2 B_2 A.$$

Les applications linéaires de l'espace vectoriel X dans lui-même sont aussi appelées *opérateurs linéaires* (agissant dans X). Pour abréger, on écrit $\mathscr{L}(X)$ au lieu de $\mathscr{L}(X, X)$. Ce n'est pas seulement un espace vectoriel, mais aussi un anneau (voire une algèbre sur K) par rapport à la composition. Ceci étant, $\mathscr{L}(X)$ est un anneau unitaire ; son élément unité est une application identique $I_X : X \to X$; $AI_X = I_X A = A$ pour tout $A \in \mathscr{L}(X)$; I_X est appelé *opérateur unitaire.*

2. Notation matricielle des applications linéaires de $\mathbf{R}^n$ dans $\mathbf{R}^m$

Il est rationnel d'écrire les vecteurs sous forme de matrices colonnes (et non pas de lignes) : pour un vecteur $x \in \mathbf{R}^n$ nous allons utiliser souvent la notation suivante :

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

où $x_1, \ldots, x_n$ sont les coordonnées de ce vecteur par rapport à la base standard $(e^1, \ldots, e^n)$ de $\mathbf{R}^n$.

Soient maintenant $A_1, \ldots, A_m$ les composantes d'une application linéaire $A : \mathbf{R}^n \to \mathbf{R}^m$, i.e. $A_i = \pi_i \circ A$ $(i = 1, \ldots, m)$, où π_i sont des fonctions de coordonnées sur $\mathbf{R}^m$, $\pi_i \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = y_i$, de sorte que

$$Ax = \begin{pmatrix} A_1 x \\ \vdots \\ A_m x \end{pmatrix} \text{ pour tout } x \in \mathbf{R}^n. \tag{1}$$

Comme π_i sont des fonctions linéaires (voir n° 5.2.2), A_i sont également des fonctions linéaires. Elles peuvent être représentées par n'importe quelles fonctions linéaires sur $\mathbf{R}^n$, i.e. quelles que soient m fonctions linéaires $A_1, \ldots, A_m$ sur $\mathbf{R}^n$, l'application $A : \mathbf{R}^n \to \mathbf{R}^m$ définie par la formule (1) est linéaire et $A_1, \ldots, A_m$ sont ses composantes. Comme toutes les fonctions linéaires sur $\mathbf{R}^n$, elles sont continues. D'où il résulte (voir n° 5.2.2, 5°) que *toute application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est continue.*

Etant une fonction linéaire, chaque A_i est donnée par ses coefficients $A_{i_1}, \ldots, A_{i_n}$:

$$A_i x = A_{i_1} x_1 + \ldots + A_{i_n} x_n \text{ pour tout } x \in \mathbf{R}^n ; \tag{2}$$

ceci étant,

$$A_{ij} = A_i e^j \ (i = 1, \ldots, m ; j = 1, \ldots, n). \tag{3}$$

Les nombres A_{ij} sont habituellement disposés en tableau

$$\begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix}$$

de m lignes et n colonnes, qu'on appelle *matrice de l'application linéaire* A (dans un couple de bases standard) ; en vertu des formules (2) et (3), la j-ième colonne de cette matrice est présentée par le vecteur Ae^j. Il résulte des formules (1) et (2) que

$$(4)\qquad Ax = \begin{pmatrix} A_1 x \\ \vdots \\ A_m x \end{pmatrix} = \begin{pmatrix} A_{11}x_1 + & \dots & + A_{1n}x_n \\ \vdots & & \vdots \\ A_{m1}x_1 + & \dots & + A_{mn}x_n \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

conformément à la règle générale de multiplication des matrices (ce qui justifie la commodité de la notation des vecteurs sous forme des colonnes). La formule (4) montre que l'application A est bien définie par sa matrice qui sera notée (A) ou (A_{ij}).

A l'addition des applications linéaires $\mathbf{R}^n \to \mathbf{R}^m$ et à la multiplication de celles-ci par des scalaires on associe les opérations analogues sur les matrices : $(A + B) = (A) + (B)$, $(\lambda A) = \lambda(A)$; d'où il découle que $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ est de dimension mn. La multiplication des matrices correspond à la composition des applications linéaires. Les matrices des opérateurs linéaires dans $\mathbf{R}^n$ sont des matrices carrées. En particulier, la matrice de l'application identique $I_{\mathbf{R}^n}$ est la matrice unité

$$(I_{\mathbf{R}^n}) = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{pmatrix}$$

qui a des unités à la diagonale et des zéros à toutes les autres positions. A chaque opérateur linéaire A dans $\mathbf{R}^n$ correspond le nombre $\det A$, *déterminant de la matrice* (A) de cet opérateur ; $\det I_{\mathbf{R}^n} = 1$. Lorsqu'on fait la composition des opérateurs linéaires A et B dans $\mathbf{R}^n$, leurs déterminants se multiplient : $\det (BA) = \det B \cdot \det A$.

3. Norme sur $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$

Dans les deux numéros précédents nous avons cité certains faits nécessaires puisés dans le cours d'algèbre. Considérons maintenant la notion de

norme d'une application linéaire et ses propriétés fondamentales, qui sont spécifiques pour l'analyse mathématique.

DÉFINITION 1. On appelle *norme d'une application linéaire* $A : \mathbf{R}^n \to \mathbf{R}^m$ la borne inférieure $\|A\|$ de l'ensemble

$$\mathrm{P}_A = \{\varrho \geqslant 0 \mid \|Ax\| \leqslant \varrho \|x\| \text{ pour tout } x \in \mathbf{R}^n\}.$$

Soit $x \in \mathbf{R}^n$. Alors $Ax = A\left(\sum_{j=1}^{n} x_j e^j\right) = \sum_{j=1}^{n} x_j A e^j$ et, par conséquent, en vertu de l'inégalité de Cauchy,

$$(5) \qquad \|Ax\| \leqslant \sum_{j=1}^{n} |x_j| \, \|Ae^j\| \leqslant \sqrt{\sum_{j=1}^{n} \|Ae^j\|^2} \sqrt{\sum_{j=1}^{n} x_j^2} = \sqrt{\sum_{j=1}^{n} \|Ae^j\|^2} \, \|x\|.$$

Comme cela est juste pour tout $x \in \mathbf{R}^n$, on a $\sqrt{\sum_{j=1}^{n} \|Ae^j\|^2} \in \mathrm{P}_A$, de sorte que P_A n'est pas vide et, par conséquent, $\|A\| < +\infty$. D'autre part, $\|A\| \geqslant 0$ puisque $\mathrm{P}_A \geqslant 0$. Ainsi donc, $\|A\|$ est un nombre réel positif. Précisons ses propriétés essentielles.

1° $\|Ax\| \leqslant \|A\| \, \|x\|$ *pour tout* $x \in \mathbf{R}^n$. En effet, comme $\|A\| = \inf \mathrm{P}_A$, il existe pour tout $\varepsilon > 0$ un $\varrho \in \mathrm{P}_A$ satisfaisant à l'inégalité $\varrho < \|A\| + \varepsilon$. Alors $\|Ax\| \leqslant \varrho \|x\| \leqslant (\|A\| + \varepsilon)\|x\| = \|A\| \, \|x\| + \varepsilon \|x\|$. En faisant tendre ε vers zéro, on obtient à la limite $\|Ax\| \leqslant \|A\| \, \|x\|$.

2° $\|A\| = 0 \Leftrightarrow A = 0$. En effet, si $\|A\| = 0$, on a $(0 \leqslant) \|A\| \leqslant \leqslant \|A\| \, \|x\| = 0$ en vertu de 1°, de sorte que $Ax = 0$ pour tout $x \in \mathbf{R}^n$, i.e. $A = 0$. Inversement : si $A = 0$, il vient $\mathrm{P}_A = [0, +\infty[$ et, par conséquent, $\|A\|$ $(= \inf \mathrm{P}_A) = 0$.

3° $\|A + B\| \leqslant \|A\| + \|B\|$. En effet, pour tout $x \in \mathbf{R}^n$ on a en vertu de 1°

$$\|(A + B)x\| = \|Ax + Bx\| \leqslant \|Ax\| + \|Bx\| \leqslant \|A\| \, \|x\| + \|B\| \, \|x\| = \\ = (\|A\| + \|B\|)\|x\|,$$

de sorte que $\|A\| + \|B\| \in \mathrm{P}_{A+B}$. Alors $\|A + B\|$ $(= \inf \mathrm{P}_{A+B}) \leqslant \|A\| + + \|B\|$.

4° $\|\lambda A\| = |\lambda| \, \|A\|$ *pour tout* $\lambda \in \mathbf{R}$. En effet, pour $\lambda = 0$ cela découle de 2°. Soit $\lambda \neq 0$. En vertu de 1°, on a pour tout $x \in \mathbf{R}^n$

$$\|(\lambda A)x\| = \|\lambda(Ax)\| = |\lambda| \, \|Ax\| \leqslant |\lambda| \, \|A\| \, \|x\|,$$

de sorte que $|\lambda| \, \|A\| \in \mathrm{P}_{\lambda A}$, d'où $\|\lambda A\| \leqslant |\lambda| \, \|A\|$. De façon analogue, $\|\lambda^{-1}(\lambda A)\| \leqslant |\lambda^{-1}| \, \|\lambda A\|$. On obtient alors

$$\|\lambda A\| \leqslant |\lambda| \, \|A\| = |\lambda| \, \|\lambda^{-1}(\lambda A)\| \leqslant |\lambda| \, |\lambda^{-1}| \, \|\lambda A\| = \|\lambda A\|$$

et, par conséquent, $\|\lambda A\| = |\lambda| \, \|A\|$.

Les propriétés 2° à 4° signifient que $\|A\|$ est une norme sur $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$.

5° Si $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ et $B \in \mathscr{L}(\mathbf{R}^m, \mathbf{R}^k)$, on a $\|BA\| \leqslant \|B\|\,\|A\|$. En effet, pour tout $x \in \mathbf{R}^n$ on obtient en vertu de 1°

$$\|BAx\| \leqslant \|B\|\,\|Ax\| \leqslant \|B\|\,\|A\|\,\|x\|,$$

de sorte que $\|B\|\,\|A\| \in \mathrm{P}_A$ et, par conséquent, $\|BA\| \leqslant \|B\|\,\|A\|$.

6° $\|I_{\mathbf{R}^n}\| = 1$. En effet,

$$\|I_{\mathbf{R}^n}\| = \inf \{\varrho \geqslant 0 \mid \|I_{\mathbf{R}^n}x\| (= \|x\|) \leqslant \varrho\|x\| \text{ pour tout } x \in \mathbf{R}^n\} = 1.$$

§ 2. Applications différentiables de $\mathbf{R}^n$ dans $\mathbf{R}^m$

1. Notion d'application différentiable de $\mathbf{R}^n$ dans $\mathbf{R}^m$

DÉFINITION 1. On dit qu'une application f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est *différentiable au point* x si D_f est un voisinage de x et l'accroissement de f en ce point peut être présenté sous la forme

$$f(x + \Delta x) - f(x) = A\Delta x + \alpha(x, \Delta x)\|x\|,$$

où A est une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$, et $\alpha(x, \Delta x) \to 0$ lorsque $\Delta x \to 0$; A est appelée *dérivée* de l'application f au point x et est notée $f'(x)$. Ainsi donc,

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + \alpha(x, \Delta x)\|x\|.$$

L'application f est dite *différentiable* si elle l'est en tout point $x \in D_f$.

En tenant compte de la définition adoptée de la limite, on obtient $\alpha(x, 0) = 0$. Si x est fixe, on écrira $\alpha(\Delta x)$ au lieu de $\alpha(x, \Delta x)$.

La définition 6.3.1 d'une fonction réelle différentiable est un cas particulier de la définition 1 (pour $m = 1$). Mais $\alpha(x, \Delta x)$ pour $m > 1$ est un vecteur (de $\mathbf{R}^m$).

Les affirmations des deux exemples qui suivent et les théorèmes 1 à 3 sont démontrés tout comme pour $m = 1$ (voir n° 6.3.1).

EXEMPLE 1. Une application constante $f : \mathbf{R}^n \to \mathbf{R}^m$ est différentiable et $f'(x) \equiv 0$.

EXEMPLE 2. Une application linéaire $A : \mathbf{R}^n \to \mathbf{R}^m$ est différentiable et $A'(x) = A$ pour tout $x \in \mathbf{R}^n$.

THÉORÈME 1. *Si les applications f et g de $\mathbf{R}^n$ dans $\mathbf{R}^m$ sont différentiables au point x, leur somme $f + g$ est différentiable en ce point et $(f + g)'(x) = f'(x) + g'(x)$.*

THÉORÈME 2. *Si une application f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est différentiable au point x, il en est de même de λf pour tout $\lambda \in \mathbf{R}$, et $(\lambda f)'(x) = \lambda f'(x)$.*

EXERCICE 1. Montrer que si deux applications f et g de $\mathbf{R}^n$ dans $\mathbf{R}^m$ sont différentiables au point x, il en est de même de l'application $\langle f, g\rangle : \mathbf{R}^n \to \mathbf{R}$ définie pour tout $z \in D_f \cap D_g$ par la formule $\langle f, g\rangle(z) = \langle f(z), g(z)\rangle$. Quelle formule définit sa dérivée ?

THÉORÈME 3. *Une application $f: \mathbf{R}^n \to \mathbf{R}^m$ différentiable au point x est continue en ce point.*

2. Composée d'applications différentiables

THÉORÈME 4. *Soit $f: \mathbf{R}^n \to \mathbf{R}^m$ une application différentiable au point x et soit $g: \mathbf{R}^m \to \mathbf{R}^l$ une application différentiable au point $y = f(x)$. La composée $g \circ f$ est alors différentiable au point x, et*

$$(g \circ f)'(x) = g'(f(x))\, f'(x) \tag{1}$$

où le second membre est la composée des applications linéaires $f'(x)$: $\mathbf{R}^n \to \mathbf{R}^m$ et $g'(f(x)) : \mathbf{R}^m \to \mathbf{R}^l$.

Démonstration. Comme $x \in D_f$ et $f(x) \in D_g$, il vient $x \in D_{g \circ f}$. Par ailleurs, puisque f est continue au point x en vertu du théorème 3 et que D_g est un voisinage du point $f(x)$, $f^{-1}(D_g)$ est un voisinage de x. Or $D_{g \circ f} = f^{-1}(D_g)$ (voir propriété 2.1.2, 5°). Ainsi donc, $D_{g \circ f}$ est un voisinage de x. Par hypothèse,

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + \alpha_1(\Delta x)\|\Delta x\| \tag{2}$$

où $\alpha_1(\Delta x) \to 0$ lorsque $\Delta x \to 0$,

$$g(y + \Delta y) - g(y) = g'(y)\Delta y + \alpha_2(\Delta y)\|\Delta y\| \tag{3}$$

où $\alpha_2(\Delta y) \to 0$ lorsque $\Delta y \to 0$. Soit maintenant $x + \Delta x \in D_{g \circ f}$. On peut alors prendre, en particulier, $\Delta y = \Delta f = f(x + \Delta x) - f(x)$ dans la formule (3) ; en effet, si $x + \Delta x \in D_{g \circ f}$, on a $x + \Delta x \in D_f$ et $y + \Delta f = f(x) + \Delta f = f(x + \Delta x) \in D_g$. Compte tenu de (2) et de la linéarité de $g'(f(x))$, on obtient

$$\begin{aligned} g(f(x + \Delta x)) - g(f(x)) &= \\ = g'(f(x))\, f'(x)\Delta x &+ g'(f(x))(\alpha_1(\Delta x))\|\Delta x\| + \alpha_2(\Delta f)\|\Delta f\| = \\ &= g'(f(x))\, f'(x)\Delta x + \alpha(\Delta x)\|\Delta x\| \end{aligned}$$

où

$$\alpha(\Delta x) = \begin{cases} g'(f(x))(\alpha_1(\Delta x)) + \alpha_2(\Delta f)\dfrac{\|\Delta f\|}{\|\Delta x\|} & \text{pour } \Delta x \neq 0, \\ 0 & \text{pour } \Delta x = 0. \end{cases} \tag{4}$$

Puisque $g'(f(x))\, f'(x)$ est une application linéaire en tant que composée d'applications linéaires, il reste à démontrer que $\alpha(\Delta x) \to 0$ quand $\Delta x \to 0$. Comme $\alpha_1(\Delta x) \to 0$ quand $\Delta x \to 0$, et $g'(f(x))$ étant une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est continue et égale à zéro au point 0, on a

$$g'(f(x))(\alpha_1(\Delta x)) \to 0 \text{ quand } \Delta x \to 0. \tag{5}$$

Comme $\alpha_2(\Delta y) \to 0$ lorsque $\Delta y \to 0$, et en vertu du théorème 3, $\Delta f \to 0$ lorsque $\Delta x \to 0$, on obtient

(6) $$\alpha_2(\Delta f) \to 0 \text{ lorsque } \Delta x \to 0.$$

D'autre part, vu qu'en vertu de (2),

$$\frac{\|\Delta f\|}{\|\Delta x\|} = \left\| f'(x) \frac{\Delta x}{\|\Delta x\|} + \alpha_1(\Delta x) \right\| \leqslant \left\| f'(x) \frac{\Delta x}{\|\Delta x\|} \right\| + \|\alpha_1(\Delta x)\| \leqslant$$
$$\leqslant \|f'(x)\| \left\| \frac{\Delta x}{\|\Delta x\|} \right\| + \|\alpha_1(\Delta x)\| = \|f'(x)\| + \|\alpha_1(\Delta x)\|,$$

et $\|\alpha_1(\Delta x)\| \to 0$ lorsque $\Delta x \to 0$, $\dfrac{\|\Delta f\|}{\|\Delta x\|}$ est borné dans un voisinage assez petit du point $\Delta x = 0$. Joint à (6), ceci montre que

(7) $$\alpha_2(\Delta f) \frac{\|\Delta f\|}{\|\Delta x\|} \to 0 \text{ lorsque } \Delta x \to 0.$$

Enfin, il résulte de (4), (5) et (7) que $\alpha(\Delta x) \to 0$ lorsque $\Delta x \to 0$.

EXEMPLE 3. Etant donné une application $g : \mathbf{R}^n \to \mathbf{R}^m$, ajoutons l nouveaux arguments en posant

$$h(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+l}) = g(x_1, \ldots, x_n)$$

pour tout $(x_1, \ldots, x_n) \in D_g$ et tout $(x_{n+1}, \ldots, x_{n+l}) \in \mathbf{R}^l$ (de sorte que $D_h = D_g \times \mathbf{R}^l$). *Si g est différentiable au point* $(x_1^{(0)}, \ldots, x_n^{(0)})$, *h est différentiable au point* $(x_1^{(0)}, \ldots, x_n^{(0)}, x_{n+1}^{(0)}, \ldots, x_{n+l}^{(0)})$ *pour toutes les valeurs* $x_{n+1}^{(0)}, \ldots, x_{n+l}^{(0)}$ *des arguments ajoutés, et*

(8) $$h'(x_1^{(0)}, \ldots, x_n^{(0)}, x_{n+1}^{(0)}, \ldots, x_{n+l}^{(0)}) = g'(x_1^{(0)}, \ldots, x_n^{(0)}) \circ \pi$$

où $\pi = \mathrm{pr}_{1\ldots n} : \mathbf{R}^{n+l} \to \mathbf{R}^n$. En effet, $h = g \circ \pi$; mais π est partout différentiable en tant qu'application linéaire, et $\pi'(x) = \pi$ pour tout $x \in \mathbf{R}^{n+l}$. Il ne reste qu'à appliquer le théorème 4. Ainsi donc, *l'adjonction de nouveaux arguments ne viole pas la différentiabilité.*

Pour $m = 1$ il résulte de la formule (8) que

(9) $$\frac{\partial h}{\partial x_j}(x_i^{(0)}, \ldots, x_{n+l}^{(0)}) = \begin{cases} \dfrac{\partial g}{\partial x_j}(x_i^{(0)}, \ldots, x_n^{(0)}) \text{ si } i \leqslant j \leqslant n, \\ 0 \text{ si } n+1 \leqslant j \leqslant n+l. \end{cases}$$

En effet, supposons que $d^1, \ldots, d^{n+l}$ soient des vecteurs de la base standard dans $\mathbf{R}^{n+l}$ et que $e^1, \ldots, e^n$ soient ceux dans $\mathbf{R}^n$. Vu que

$$\pi(d^j) = \mathrm{pr}_{i\ldots n} d^j = \begin{cases} e^j & \text{si } 1 \leqslant j \leqslant n, \\ 0 & \text{si } n+1 \leqslant j \leqslant n+l, \end{cases}$$

on a

$$\frac{\partial h}{\partial x_j}(x_1^{(0)}, \ldots, x_{n+j}^{(0)}) = h'(x_1^{(0)}, \ldots, x_{n+l}^{(0)})d^j = g'(x_1^{(0)}, \ldots, x_n^{(0)})\pi(d^j) =$$

$$= \begin{cases} g'(x_1^{(0)}, \ldots, x_n^{(0)})e^j = \dfrac{\partial d}{\partial x_j}(x_1^{(0)}, \ldots, x_n^{(0)}) & \text{si } 1 \leqslant j \leqslant n, \\ 0 & \text{si } n+1 \leqslant j \leqslant n+l. \end{cases}$$

EXEMPLE 4 (propriété locale de la dérivée). *Si une application f: $\mathbf{R}^n \to \mathbf{R}^m$ est différentiable en un point a et U est un voisinage de a contenu dans D_f, la restriction $f|_U$ de f à U est différentiable en a et*

$$(f|_U)'(a) = f'(a).$$

En effet, $f|_U = f \circ i_U$, où i_U est la restriction de $I_{\mathbf{R}^n}$ à U. Etant donné que $D_{i_U} = U$ est un voisinage du point a et $i_U(x) - i_U(a) = x - a = I_{\mathbf{R}^n}(x-a) + 0 \cdot \|x - a\|$ pour tout $x \in D_{i_U}$, i_U est différentiable en a et $i_U'(a) = I_{\mathbf{R}^n}$. Vu que f est différentiable au point $i_U(a)$ $(=a)$, le théorème 4 dit que $f|_U$ est différentiable au point a et $(f|_U)'(a) = f'(i_U(a))i_U'(a) = f'(a)I_{\mathbf{R}^n} = f'(a)$.

3. Matrice jacobienne

THÉORÈME 5. *Une application f de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est différentiable au point x si et seulement si ses composantes $f_1, \ldots, f_m$ sont différentiables en ce point. La dérivée de chaque composante est une composante correspondante de la dérivée de l'application f.*

Démonstration. En vertu de la définition 5.1.2, $f_k = \pi_k \circ f$ où π_k est la k-ième fonction de coordonnées sur $\mathbf{R}^m$. Etant linéaire (voir n° 6.2.2), π_k est partout différentiable (voir exemple 6.3.2). Par conséquent, la différentiabilité de f au point x entraîne celle de ses composantes f_k d'après le théorème 4. D'autre part, $f = \sum_{k=1}^{n} i_k \circ f_k$ (voir formule 5.1(3)) où i_k est la k-ième injection de $\mathbf{R}$ dans $\mathbf{R}^m$ (voir n° 5.1.2). Etant linéaires (voir n° 1 du § 1), i_k sont partout différentiables. Par conséquent, la différentiabilité de toutes les composantes f_k entraîne celle de f d'après les théorèmes 4 et 1. Enfin, si f est différentiable au point x, on obtient $f_k'(x) = (\pi_k \circ f)'(x) = \pi_k'(f(x)) \cdot f'(x)$ (voir théorème 4) $= \pi_k f'(x)$ (voir exemple 6.3.2) $= [f'(x)]_k$.

THÉORÈME 6. *Si une application* $f: \mathbf{R}^n \to \mathbf{R}^m$ *de composantes* $f_1, \ldots$ $\ldots, f_n$ *est différentiable au point* x, *sa dérivée* $f'(x)$ *possède dans les bases standard* (voir n° 1.1.1) *la matrice suivante* :

$$(*) \qquad (f'(x)) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \ldots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Démonstration. Les lignes de la matrice de l'application linéaire sont les matrices de ses composantes. En vertu du théorème 5, on a $[f'(x)]_k =$ $= f_k'(x)$, et en vertu de la formule 6.3(4), $(f_k'(x)) = \left(\frac{\partial f_k}{\partial x_1}(x), \ldots\right.$ $\left.\ldots, \frac{\partial f_k}{\partial x_n}(x)\right)$.

REMARQUE. Vu qu'une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ est bien définie par sa matrice, il résulte en particulier du théorème 6 que *la dérivée* $f'(x)$ *de l'application* f *au point* x *est univoquement définie.*

DÉFINITION 2. La matrice $(*)$ est appelée *matrice jacobienne* de l'application f au point x. Pour $m = n$, le déterminant $\det f'(x)$ de cette matrice est appelé *jacobien* de l'application f au point x et noté $\frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}(x)$.

Si $m = 1$, la matrice jacobienne $(f'(x))$ est représentée par la ligne $\left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)$ des coordonnées du gradient. Si de plus $n = 1$, la matrice jacobienne $(f'(x))$ est formée par un seul nombre $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$.

THÉORÈME 7. *Si les composantes* $f_1, \ldots, f_m$ *de l'application* f: $\mathbf{R}^n \to \mathbf{R}^m$ *ont, dans un voisinage du point* x, *des dérivées partielles par rapport à tous les arguments, qui sont continues en* x, *l'application* f *est différentiable en ce point.*

Démonstration. En vertu du théorème 6.3.6, les fonctions $f_1, \ldots, f_m$ sont différentiables au point x, d'où il résulte en vertu du théorème 5 que f est différentiable en x.

EXEMPLE 5. Soit f une application de $\mathbf{R}^2$ dans $\mathbf{R}^2$ ayant les composantes $f_1(r, \varphi) = r\cos\varphi$, $f_2(r, \varphi) = r\sin\varphi$. Elles ont partout des dérivées partielles continues

$$(\partial_1 f_1)(r, \varphi) = \cos\varphi, \quad (\partial_2 f_1)(r, \varphi) = -r\sin\varphi,$$
$$(\partial_1 f_2)(r, \varphi) = \sin\varphi, \quad (\partial_2 f_2)(r, \varphi) = r\cos\varphi,$$

de sorte que f est différentiable. Son jacobien est

$$\frac{D(f_1, f_2)}{D(r, \varphi)}(r, \varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

THÉORÈME 8. *Soit $f : \mathbf{R}^n \to \mathbf{R}^m$ une application différentiable au point x et soit $g : \mathbf{R}^m \to \mathbf{R}^l$ une application différentiable au point $y = f(x)$, de sorte que $h = g \circ f$ est différentiable en x en vertu du théorème 4. Alors*

$$\frac{\partial h_i}{\partial x_j}(x) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(y)\frac{\partial f_k}{\partial x_j}(x) \ (i = 1, \ldots, l;\ j = 1, \ldots, n). \tag{10}$$

Démonstration. D'après le théorème 4, $h'(x) = g'(y)f'(x)$. Par suite, $(h'(x)) = (g'(y))(f'(x))$, i.e. en vertu du théorème 6

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x) & \ldots & \frac{\partial h_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial h_l}{\partial x_1}(x) & \ldots & \frac{\partial h_l}{\partial x_n}(x) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(y) & \ldots & \frac{\partial g_1}{\partial y_m}(y) \\ \vdots & & \vdots \\ \frac{\partial g_l}{\partial y_1}(y) & \ldots & \frac{\partial g_l}{\partial y_m}(y) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \ldots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Selon la règle de multiplication des matrices, l'élément $\frac{\partial h_i}{\partial x_j}(x)$ de la matrice du premier membre, qui se trouve à intersection de la i-ième ligne et de la j-ième colonne, est égal au « produit scalaire » de la i-ième ligne de la première matrice dans le second membre par la j-ième colonne de la deuxième, ce qui donne la formule (10).

4. Dérivation des fonctions composées

THÉORÈME 9. *Supposons que les arguments $y_1, \ldots, y_m$ de la fonction $g : \mathbf{R}^m \to \mathbf{R}$ soient remplacés par des fonctions $f_i : \mathbf{R}^n \to \mathbf{R}$ $(i = 1, \ldots, m)$ ayant un ensemble de définition commun. Si ces fonctions sont différentiables au point $y^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, et g, au point correspondant $y^{(0)}(f_1(x^{(0)}, \ldots, f_m(x^{(0)}))$, la fonction*

$$h(x) = g(f_1(x), \ldots, f_m(x))$$

est différentiable au point $x^{(0)}$ *et*

$$(11)\qquad \frac{\partial h}{\partial x_j}(x^{(0)}) = \sum_{i=1}^{m} \frac{\partial g}{\partial y_i}(y^{(0)}) \frac{\partial f_i}{\partial x_j}(x^{(0)}) \quad (j = 1, \dots, n).$$

Démonstration. Soit f une application de $\mathbf{R}^n$ dans $\mathbf{R}^m$ ayant des composantes $f_1, \dots, f_m$. On a $y^{(0)} = f(x^{(0)})$ et $h = g \circ f$. Comme g est différentiable au point $y^{(0)}$ et la différentiabilité des fonctions f_i au point $x^{(0)}$ entraîne, d'après le théorème 5, la différentiabilité de l'application f, il reste à appliquer les théorèmes 5 et 8.

Nous savons que le théorème 9 ne représente pas le cas le plus général de formation de la fonction composée ; chacune des fonctions substituées f_i dépend en général de son propre groupe des arguments. Mais le cas où ces groupes sont distincts se réduit au cas envisagé dans le théorème 9 par adjonction des arguments manquants. Soit, par exemple,

$$f(x, y, z, t) = g(\varphi(x, y), \psi(y, z, t))$$

(voir n°s 5.1.5 et 5.2.2, 8°). *Si* φ *est différentiable au point* (x_0, y_0), ψ, *au point* (y_0, z_0, t_0) *et* g, *au point* (u_0, v_0) *où* $u_0 = \varphi(x_0, y_0)$ *et* $v_0 = \psi(y_0, z_0, t_0)$, *la fonction* f *est différentiable au point* (x_0, y_0, z_0, t_0). En effet,

$$f(x, y, z, t) = g(\bar{\varphi}(x, y, z, t), \bar{\psi}(x, y, z, t))$$

où $\bar{\varphi}(x, y, z, t) = \varphi(x, y)$, $\bar{\psi}(x, y, z, t) = \psi(y, z, t)$. Comme φ est différentiable au point (x_0, y_0), $\bar{\varphi}$ est différentiable au point (x_0, y_0, z_0, t_0) en vertu de l'exemple 3. De façon analogue, la différentiabilité de ψ au point (y_0, z_0, t_0) entraîne celle de $\bar{\psi}$ au point (x_0, y_0, z_0, t_0). Le théorème 9 dit alors que f est différentiable en ce point. Ceci étant, on obtient d'après les formules (11) et (9),

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0, t_0) = \frac{\partial g}{\partial u}(u_0, v_0) \frac{\partial \bar{\varphi}}{\partial x}(x_0, y_0, z_0, t_0) + \frac{\partial g}{\partial v}(u_0, v_0) \frac{\partial \bar{\psi}}{\partial x}(x_0, y_0, z_0, t_0) = \frac{\partial g}{\partial u}(u_0, v_0) \frac{\partial \varphi}{\partial x}(x_0, y_0),$$

$$\frac{\partial f}{\partial y}(x_0, y_0, z_0, t_0) = \frac{\partial g}{\partial u}(u_0, v_0) \frac{\partial \varphi}{\partial y}(x_0, y_0) + \frac{\partial g}{\partial v}(u_0, v_0) \frac{\partial \psi}{\partial y}(y_0, z_0, t_0),$$

$$\frac{\partial f}{\partial z}(x_0, y_0, z_0, t_0) = \frac{\partial g}{\partial v}(u_0, v_0) \frac{\partial \psi}{\partial z}(y_0, z_0, t_0),$$

$$\frac{\partial f}{\partial t}(x_0, y_0, z_0, t_0) = \frac{\partial g}{\partial v}(u_0, v_0) \frac{\partial \psi}{\partial t}(y_0, z_0, t_0).$$

Notons un cas particulier important du théorème 9 lorsque $n = 1$, i.e. lorsque $f_1, \ldots, f_m$ sont des fonctions d'une seule variable réelle. La formule (11) prend dans ce cas la forme suivante

$$\frac{\mathrm{d}h}{\mathrm{d}t}(t_0) = \sum_{i=1}^{m} \frac{\partial g}{\partial y_i}(y_0)\frac{\mathrm{d}f_i}{\mathrm{d}t}(t_0). \tag{12}$$

EXEMPLE 6 (formule de dérivation d'une expression exponentielle). Soit $h(t) = \varphi(t)^{\psi(t)}$ où φ et ψ sont des fonctions dérivables au point t_0. La fonction $h(t)$ s'obtient de la fonction $g(u, v) = u^v$ par substitution $u = \varphi(t)$, $v = \psi(t)$; g est différentiable et $\frac{\partial g}{\partial u}(u, v) = vu^{v-1}$, $\frac{\partial g}{\partial v}(u, v) =$ $= u^v \ln u$. Par conséquent, h est différentiable au point t_0 d'après le théorème 9 et donc, d'après la formule (12),

$$\frac{\mathrm{d}h(t)}{\mathrm{d}t}(t_0) = \psi(t_0)\varphi(t_0)^{\psi(t_0)-1}\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t_0) + \varphi(t_0)^{\psi(t_0)}\ln\varphi(t_0)\frac{\mathrm{d}\psi}{\mathrm{d}t}(t_0).$$

EXERCICE 2. Déduire les formules de dérivation pour les fonctions $\varphi(t) + \psi(t)$, $\varphi(t)\cdot\psi(t)$ et $\dfrac{\varphi(t)}{\psi(t)}$ de $\mathbf{R}$ dans $\mathbf{R}$ où φ et ψ sont dérivables au point t_0.

5. Dérivation d'une fonction vectorielle

Les applications de $\mathbf{R}$ dans $\mathbf{R}^m$ sont appelées *fonctions vectorielles* à valeurs dans $\mathbf{R}^m$. Soit A une fonction vectorielle linéaire à valeurs dans $\mathbf{R}^m$, i.e. une application linéaire de $\mathbf{R}$ dans $\mathbf{R}^m$; A est parfaitement définie par le vecteur $A(1)$. En effet, on a $A(h) = A(h\cdot 1) = hA(1)$ pour tout $h \in \mathbf{R}$. Le vecteur $A(1)$ sera noté $\vec{A}$. Il vient donc

$$A(h) = h\vec{A} \text{ pour tout } h \in \mathbf{R}. \tag{13}$$

On voit aisément que $A \mapsto \vec{A}$ définit un isomorphisme de l'espace vectoriel $\mathscr{L}(\mathbf{R}, \mathbf{R}^m)$ sur $\mathbf{R}^m$. Ceci étant, $\|A\| = \|\vec{A}\|$ car $\|Ah\| = |h|\,\|\vec{A}\|$ pour tout $h \in \mathbf{R}$. En se fondant sur cette affirmation, on identifie souvent la fonction vectorielle linéaire A avec le vecteur correspondant $\vec{A}$.

THÉORÈME 10. *Si une fonction vectorielle f à valeurs dans $\mathbf{R}^m$ est différentiable au point t, on a*

$$f'(t) = \lim_{\Delta t \to 0}\frac{f(t + \Delta t) - f(t)}{\Delta t}. \tag{14}$$

Inversement : si $t \in]D_f[$ et $\dfrac{f(t + \Delta t) - f(t)}{\Delta t}$ tend vers le vecteur $a \in \mathbf{R}^m$ lorsque $\Delta t \to 0$, la fonction f est différentiable au point t et $a = \overrightarrow{f'(t)}$.

Démonstration. Si f est différentiable au point t, on obtient, en vertu de la formule (13),

$$f(t + \Delta t) - f(t) = \Delta t \cdot \overrightarrow{f'(t)} + \alpha(t, \Delta t)|\Delta t|$$

où $\alpha(t, \Delta t) \to 0$ lorsque $\Delta t \to 0$, d'où

$$\frac{f(t + \Delta t) - f(t)}{\Delta t} = \overrightarrow{f'(t)} + \alpha(t, \Delta t)\frac{|\Delta t|}{\Delta t} \to \overrightarrow{f'(t)} \text{ lorsque } \Delta t \to 0.$$

Inversement : soit $t \in]D_f[$ et $\dfrac{f(t + \Delta t) - f(t)}{\Delta t}$ tend vers le vecteur $a \in \mathbf{R}^m$ lorsque $\Delta t \to 0$. Posons

$$\alpha(t, \Delta t) = \begin{cases} \dfrac{f(t + \Delta t) - f(t)}{\Delta t} - a & \text{pour } \Delta t \neq 0, \\ 0 & \text{pour } \Delta t = 0. \end{cases}$$

Alors $\alpha(t, \Delta t) \to 0$ lorsque $\Delta t \to 0$, $f(t + \Delta t) - f(t) = \Delta t \cdot a + \alpha(t, \Delta t)\Delta t$, et comme $h \cdot a$ est une fonction vectorielle linéaire, f est différentiable au point t et $a = \overrightarrow{f'(t)}$.

D'ordinaire $\overrightarrow{f'(t)}$ est identifiée avec $f'(t)$ et la formule

$$f'(t) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

est prise pour définition de la dérivée de la fonction vectorielle f au point t.

Il résulte de la formule (14) que $\overrightarrow{f'(t)} = \begin{pmatrix} \dfrac{df_1}{dt}(t) \\ \cdots\cdots \\ \dfrac{df_m}{dt}(t) \end{pmatrix}$ où $f_1, \ldots, f_m$ sont les composantes de la fonction vectorielle f. En effet, soit $\overrightarrow{f'(t)}_k$ la k-ième coordonnée du vecteur $\overrightarrow{f'(t)}$, i.e. $\pi_k(\overrightarrow{f'(t)})$ où π_k est la k-ième fonction de coordonnées sur $\mathbf{R}^m$. Comme π_k est linéaire et continue, on peut écrire en vertu de la formule (14) :

$$\overrightarrow{f'(t)}_k = \pi_k\left(\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}\right) = \lim_{\Delta t \to 0} \pi_k\left(\frac{f(t + \Delta t) - f(t)}{\Delta t}\right) =$$

$$= \lim_{\Delta t \to 0} \frac{\pi_k(f(t + \Delta t)) - \pi_k(f(t))}{\Delta t} = \lim_{\Delta t \to 0} \frac{f_k(t + \Delta t) - f_k(t)}{\Delta t} = \frac{df_k}{dt}(t).$$

Soit f une fonction vectorielle à valeurs dans $\mathbf{R}^m$, définie sur un intervalle non dégénéré I. On appelle *hodographe* de f la famille des points $(f(t))_{t \in I}$ de $\mathbf{R}^m$. Soit maintenant $m = 2$ ou 3 et supposons que f soit con-

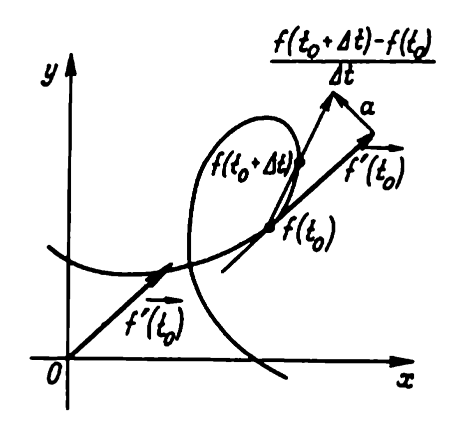


Fig. 20

tinue en un point t_0 et qu'il existe un voisinage U de ce point tel que $f(t) \neq f(t_0)$ pour tout $t \in U \cap I$ distinct de t_0. On appelle *tangente* à l'hodographe au point $M_0 = f(t_0)$ une droite L, passant par M_0, telle que la distance MN du point $M = f(t)$ à cette droite est infiniment petite devant M_0M lorsque $t \to t_0$. On voit aisément que pour cela il est nécessaire et suffisant que l'angle entre L et la sécante à l'hodographe passant par les points $f(t_0)$ et $f(t_0 + \Delta t)$ où $\Delta t \neq 0$, tende vers zéro quand $\Delta t \to 0$.

Si f est dérivable au point t_0 et $f'(t_0) \neq 0$, l'hodographe de la fonction vectorielle f a une tangente au point $f(t_0)$, et cette tangente est parallèle au vecteur $\overrightarrow{f'(t_0)}$ (fig. 20). En effet, étant dérivable au point t_0, f est continue en ce point. Par ailleurs, comme $f'(t_0)(h) = h\overrightarrow{f'(t_0)}$ et $f'(t_0) \neq 0$, on obtient $\overrightarrow{f'(t_0)} \neq 0$ et il résulte alors de la formule (14) que $f(t) \neq f(t_0)$ pour tout $t \in I \setminus \{t_0\}$ suffisamment proche de t_0. Enfin, le vecteur $\frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$ issu du point $f(t_0)$ est situé sur la sécante de l'hodographe passant par les points $f(t_0)$ et $f(t_0 + \Delta t)$, et comme $\frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t} = \overrightarrow{f'(t_0)} + \alpha$ où $\alpha \to 0$ lorsque $\Delta t \to 0$, on a

$$\cos\left(\overrightarrow{f'(t_0)}, \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}\right) = \frac{\langle \overrightarrow{f'(t_0)}, \overrightarrow{f'(t_0)} + \alpha \rangle}{\|\overrightarrow{f'(t_0)}\| \|f'(t_0) + \alpha\|} \to$$

$$\to \frac{\langle \overrightarrow{f'(t_0)}, \overrightarrow{f'(t_0)} \rangle}{\|f'(t_0)\|^2}$$

lorsque $\Delta t \to 0$, i.e. l'angle formé par la sécante et le vecteur $\overrightarrow{f'(t_0)}$ tend vers zéro.

Si l'argument d'une fonction vectorielle est le temps, l'hodographe de celle-ci est la trajectoire du point qui se meut suivant la loi $s = f(t)$ (où s est le rayon vecteur du point), et $\overrightarrow{f'(t)}$ est, comme on le voit de la formule (14), le vecteur vitesse du point à l'instant t.

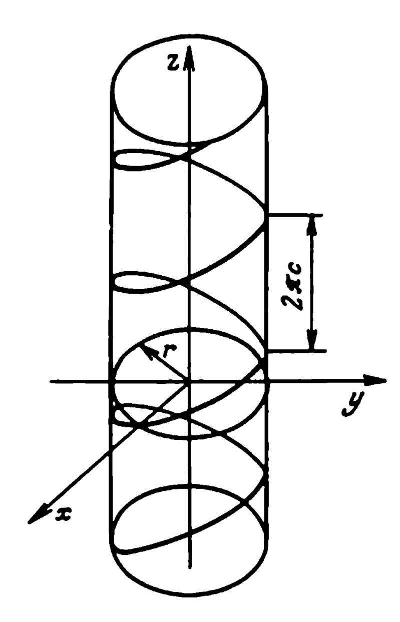


Fig. 21

EXEMPLE 7. Soit f une application de $\mathbf{R}$ dans $\mathbf{R}^3$ de composantes

$$\varphi(t) = r \cos t, \qquad \psi(t) = r \sin t, \qquad \chi(t) = ct$$

où r et c sont des constantes non nulles. Comme φ, ψ et χ sont dérivables, f est une fonction vectorielle dérivable et

$$\overrightarrow{f'(t)} = \begin{pmatrix} -r \sin t \\ r \cos t \\ c \end{pmatrix}.$$

L'hodographe de cette fonction vectorielle est une hélice située sur le cylindre circulaire droit $x^2 + y^2 = r^2$ (fig. 21).

CHAPITRE 8

DÉRIVÉES PARTIELLES ET DIFFÉRENTIELLES D'ORDRES SUPÉRIEURS

§ 1. Dérivées partielles d'ordres supérieurs

1. Dérivées partielles d'ordre 2

Soit f une fonction réelle de n variables réelles. Sa dérivée partielle par rapport à x_i, $\partial_i f = \dfrac{\partial f}{\partial x_i} = f'_{x_i}$, est aussi une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$, ayant peut-être un ensemble de définition plus étroit, soit quand même non vide. Supposons que $\partial_i f$ possède une dérivée partielle par rapport à x_j, $\partial_j(\partial_i f)$, définie sur un ensemble non vide ; on la note en général $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ ou $f''_{x_i x_j}$ et on l'appelle *dérivée partielle d'ordre* 2 *de la fonction f par rapport à x_i et x_j*. Conformément à cela, on dit que $\partial_1 f, \ldots, \partial_n f$ sont les *dérivées partielles d'ordre* 1. Ainsi donc, $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial\left(\dfrac{\partial f}{\partial x_i}\right)}{\partial x_j}$. Si $j = i$, on écrit $\dfrac{\partial^2 f}{\partial x_i^2}$ ou $f''_{x_i^2}$ au lieu de $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$. Lorsque $j \neq i$, on dit que $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ est une dérivée partielle *mixte* (par rapport à x_i et x_j). Ainsi, $f(x, y)$ peut avoir deux dérivées partielles d'ordre 1 et quatre dérivées partielles d'ordre 2, parmi celles-ci deux dérivées mixtes : f''_{xy} et f''_{yx}. Cependant, ces dernières coïncident habituellement. Par exemple, si $f(x, y) = \ln(x^2 + y^2)$, on a

$$f'_x(x, y) = \frac{2x}{x^2 + y^3}, \quad f''_{xy} = -\frac{2x \cdot 3y^2}{(x^2 + y^3)^2},$$

$$f'_y(x, y) = \frac{3x^2}{x^2 + y^3}, \quad f''_{yx} = -\frac{3y^2 \cdot 2x}{(x^2 + y^3)^2},$$

de sorte que $f''_{xy} = f''_{yx}$.

EXERCICE 1. Soit $f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$ (où a, b, c, d sont des constantes). Trouver f''_{xy} et f''_{yx}.

2. Limites répétées

Analysons le mécanisme qui est à la base de la coïncidence de f''_{xy} et

f''_{yx} (ou de leur non-coïncidence éventuelle). D'après la définition,

$$f''_{xy}(a, b) = (f'_x)'_y(a, b) = \lim_{k \to 0} \frac{f'_x(a, b+k) - f'_x(a, b)}{k} =$$

$$= \lim_{k \to 0} \frac{\lim\limits_{h \to 0} \dfrac{f(a+h, b+k) - f(a, b+k)}{h} - \lim\limits_{h \to 0} \dfrac{f(a+h, b) - f(a, b)}{h}}{k} =$$

$$= \lim_{h \to 0} \lim_{k \to 0} \frac{f(a+h, b+k) - f(a, b+k) - f(a+h, b) + f(a, b)}{kh}.$$

De façon analogue,

$$f''_{yx}(a, b) =$$

$$= \lim_{k \to 0} \lim_{h \to 0} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk}.$$

Ainsi donc, *les dérivées mixtes* $f''_{xy}(a, b)$ *et* $f''_{yx}(a, b)$ *sont les limites répétées d'une même fonction* $\dfrac{\Delta(h, k)}{hk}$ *où*

$$\Delta(h, k) = f(a+h, b+k) - f(a, b+k) - f(a+h, b) + f(a, b),$$

limites qui ne se distinguent que par l'ordre de $h \to 0$ *et* $k \to 0$. Quelquefois cette distinction est sensible. Ainsi,

$$\lim_{k \to 0} \lim_{h \to 0} \frac{h^2 - k^2}{h^2 + k^2} = -1 \; ; \qquad \lim_{h \to 0} \lim_{k \to 0} \frac{h^2 - k^2}{h^2 + k^2} = 1.$$

C'est pourquoi, pour la fonction

$$f(x, y) = \begin{cases} xy \dfrac{x^2 - y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0), \end{cases}$$

on a

$$f''_{xy}(0, 0) = \lim_{k \to 0} \lim_{h \to 0} \frac{f(h, k) - f(0, k) - f(h, 0) + f(0, 0)}{kh} =$$

$$= \lim_{k \to 0} \lim_{h \to 0} \frac{f(h, k)}{hk} = -1$$

et

$$f''_{yx}(0, 0) = \lim_{h \to 0} \lim_{k \to 0} \frac{f(h, k)}{hk} = 1,$$

de sorte que $f''_{xy}(0, 0) \neq f''_{yx}(0, 0)$.

Cependant, sous certaines conditions les limites répétées existent et sont égales.

LEMME. *Soit $g(u, v)$ une fonction définie pour les seuls points (u, v) du voisinage cubique $Q((u_0, v_0)\,;\,\eta)$ du point (u_0, v_0), qui ont $u \neq u_0$ et $v \neq v_0$. Supposons qu'il existe les limites finies :*

$$\lim_{(u,v)\to(u_0,v_0)} g(u, v),$$

$$\lim_{u\to u_0} g(u, v) \text{ pour tout } v \text{ fixe tel que } 0 < |v - v_0| < \eta,$$

$$\lim_{v\to v_0} g(u, v) \text{ pour tout } u \text{ fixe tel que } 0 < |u - u_0| < \eta.$$

Les limites répétées $\lim\limits_{v\to v_0} \lim\limits_{u\to u_0} g(u, v)$ *et* $\lim\limits_{u\to u_0} \lim\limits_{v\to v_0} g(u, v)$ *existent et sont égales à* $\lim\limits_{(u,v)\to(u_0,v_0)} g(u, v)$.

Démonstration. Soit $l = \lim\limits_{(u,v)\to(u_0,v_0)} g(u, v)$. Pour tout $\varepsilon > 0$ il existe un $\delta \in]0, \eta[$ tel que

$$|g(u, v) - l| < \varepsilon \text{ pour tout } (u, v) \in D_g \cap Q((u_0, v_0)\,;\,\eta).$$

En fixant v et en passant à la limite lorsque $u \to u_0$, on obtient $\left|\lim\limits_{u\to u_0} g(u, v) - l\right| \leqslant \varepsilon$ pour tout v tel que $0 < |v - v_0| < \delta$. Mais cela signifie que $\lim\limits_{u\to u_0} g(u, v) \to l$ lorsque $v \to v_0$, i.e. $\lim\limits_{v\to v_0} \lim\limits_{u\to u_0} g(u, v)$ existe et est égal à l. La démonstration pour la deuxième limite est analogue.

3. Conditions suffisantes de l'égalité des dérivées mixtes

THÉORÈME 1. *Si la fonction $f(x, y)$ a les dérivées partielles f'_x, f'_y et f''_{xy} dans le voisinage U du point (a, b) et si f''_{xy} est continue au point (a, b), la dérivée $f''_{yx}(a, b)$ existe et est égale à $f''_{xy}(a, b)$.*

Démonstration. Il existe un $\eta > 0$ tel que $U \supset Q((a, b)\,;\,\eta)$. La fonction

$$g(h, k) = \frac{\Delta(h, k)}{hk}$$

(où $\Delta(h, k)$ est introduit au n° 2) est alors définie en tout point $(h, k) \in \in Q((0, 0)\,;\,\eta)$ tel que $h \neq 0$ et $k \neq 0$. Montrons que g satisfait aux hypothèses du lemme pour $(u_0, v_0) = (0, 0)$:

1) en fixant un k tel que $|k| \in]0, \eta[$, posons

$$\varphi(x) = f(x, b + k) - f(x, b), \text{ avec } D_\varphi =]a - \eta, a + \eta[.$$

De l'existence de f'_x dans $Q((a, b)\,;\,\eta)$ il résulte que φ est dérivable et

$$\varphi'(x) = f'_x(x, b + k) - f'_x(x, b).$$

Or $\Delta(h, k) = \varphi(a + h) - \varphi(a)$ pour tout h tel que $|h| < \eta$. Par conséquent, en vertu du théorème de Lagrange, on obtient

$$g(h, k) = \frac{\varphi(a + h) - \varphi(a)}{hk} = \frac{\varphi'(a + \theta h)}{k} =$$
$$= \frac{f'_x(a + \theta h, b + k) - f'_x(a + \theta h, b)}{k},$$

où $\theta \in]0, 1[$. Puisque le segment qui joint les points $(a + \theta h, b)$ et $(a + \theta h, b + k)$ est contenu dans $Q((a, b) ; \eta)$ et que la dérivée f''_{xy} existe dans ce voisinage, on a toujours d'après le théorème de Lagrange :

$$g(h, k) = f''_{xy}(a + \theta h, b + \theta' k)$$

où $\theta' \in]0, 1[$. Etant donné que f''_{xy} est continue au point (a, b), on en conclut que $\lim_{(h,k)\to(0,0)} g(h, k)$ existe et est égale à $f''_{xy}(a, b)$;

2) $g(h, k) = \dfrac{\varphi(a + h) - \varphi(a)}{hk} \to \dfrac{\varphi'(a)}{k} = \dfrac{f''_x(a, b + k) - f'_x(a, b)}{k}$

lorsque $h \to 0$. Ainsi donc, $\lim_{h\to 0} g(h, k)$ existe pour tout k fixe tel que $|k| \in]0, \eta[$;

3) d'autre part,

$$g(h, k) = \frac{\dfrac{f(a + h, b + k) - f(a + h, b)}{k} - \dfrac{f(a, b + k) - f(a, b)}{k}}{h} \to$$
$$\to \frac{f'_y(a + h, b) - f'_y(a, b)}{h} \text{ lorsque } k \to 0,$$

de sorte que $\lim_{k\to 0} g(h, k)$ existe et est égale à $\dfrac{f'_y(a + h, b) - f'_y(a, b)}{h}$ pour tout h fixe tel que $|h| \in]0, \eta[$.

Donc, les hypothèses du lemme sont satisfaites. Il en résulte que $\lim_{h\to 0} \lim_{k\to 0} g(h, k)$ existe et est égale à

$$\frac{f'_y(a + h, b) - f'_y(a, b)}{h},$$

i.e. compte tenu des points 3) et 1),

$$f''_{yx}(a, b) = \lim_{h\to 0} \frac{f'_y(a + h, b) - f'_y(a, b)}{h}$$

existe et est égale à $f''_{xy}(a, b)$.

REMARQUE. Au lieu de l'existence de f''_{xy} dans un voisinage du point (a, b) et de la continuité en ce point, on pourrait admettre la même chose pour f''_{yx} et démontrer que $f''_{xy}(a, b)$ existe et est égale à $f''_{yx}(a, b)$. Il faudrait alors analyser la fonction $\psi(y) = f(a + h, y) - f(a, y)$, avec h fixe, au lieu de la fonction $\varphi(x) = f(a, b + k) - f(x, b)$, avec k fixe.

THÉORÈME 2. *Si la fonction* $f(x, y)$ *possède au point* (a, b) *des dérivées partielles différentiables d'ordre* 1, *les dérivées mixtes* f''_{xy} *et* f''_{yx} *coïncident en ce point.*

Démonstration. Suivant la définition de la différentiabilité, f'_x et f'_y, et à plus forte raison, f, existent au voisinage du point (a, b) et, par conséquent, dans un voisinage de la forme $Q((a, b)\,;\,\eta)$. Posons

$$\Delta(h) = f(a + h, b + h) - f(a, b + h) - f(a + h, b) + f(a, b),$$

avec $D_\Delta =]-\eta, \eta[$, et

$$\varphi(x) = f(x, b + h) - f(x, b),$$

avec $D_\varphi =]a - \eta, a + \eta[$, de sorte que $\Delta(h) = \varphi(a + h) - \varphi(a)$. Il résulte de l'existence de f'_x que φ est dérivable et $\varphi'(x) = f'_x(x, b + h) - f'_x(x, b)$. Par conséquent, on a suivant le théorème de Lagrange :

$$(1) \quad \Delta(h) = \varphi'(a + \theta_1 h)h = [f'_x(a + \theta_1 h, b + h) - f'_x(a + \theta_1 h, b)]h$$

où $\theta_1 \in]0, 1[$. Par hypothèse, f'_x est différentiable au point (a, b), i.e.

$$f'_x(x, y) = f'_x(a, b) + f''_{x^2}(a, b)(x - a) + f''_{xy}(a, b)(y - b) + \varepsilon_1(x - a, y - b)\varrho$$

où $\varrho = \sqrt{(x - a)^2 + (y - b)^2}$ et $\varepsilon_1 \to 0$ lorsque $(x, y) \to (a, b)$. En particulier,

$$f'_x(a + \theta_1 h, b + h) = f'_x(a, b) + f''_{x^2}(a, b)\theta_1 h + f''_{xy}(a, b)h + \varepsilon_1(\theta_1 h, h)\sqrt{1 + \theta_1^2}\,|h|\,;$$

$$f'_x(a + \theta_1 h, b) = f'_x(a + b) + f''_{x^2}(a, b)\theta_1 h + \varepsilon_1(\theta_1 h, 0)\theta_1 |h|.$$

En retranchant la deuxième expression de la première et en tenant compte de la formule (1), on obtient

$$\Delta(h) = f''_{xy}(a, b)h^2 + \alpha(h)h^2$$

où $\alpha(h) = \pm[\varepsilon_1(\theta_1 h, h)\cdot\sqrt{1 + \theta_1^2} - \varepsilon_1(\theta_1 h, 0)\theta_1]$. Comme $\alpha(h) \to 0$ lorsque $h \to 0$, on en conclut que

$$(2) \qquad f''_{xy}(a, b) = \lim_{h \to 0} \frac{\Delta(h)}{h^2}.$$

D'autre part, en fixant h, mettons $\psi(y) = f(a + h, y) - f(a, y)$, avec $D_\psi =]a - \eta, a + \eta[$. Il résulte de l'existence de f'_y que ψ est dérivable et

$\psi'(y) = f'_y(a + h, y) - f'_y(a, y)$. Par conséquent,

$$(3) \quad \Delta(h) = \psi(b + h) - \psi(b) = \psi'(b + \theta_2 h)h = \\ = [f'_y(a + h, b + \theta_2 h) - f'_y(a, b + \theta_2 h)]h,$$

où $\theta_2 \in]0, 1[$. Or f'_y est différentiable au point (a, b), i.e.

$$f'_y(x, y) = f'_y(a, b) + f''_{yx}(a, b)(x - a) + f''_{y^2}(a, b)(y - b) + \\ + \varepsilon_2(x - a, y - b)\varrho$$

où $\varepsilon_2(x - a, y - b) \to 0$ lorsque $(x, y) \to (a, b)$. En particulier,

$$f'_y(a + h, b + \theta_2 h) = f'_y(a, b) + f''_{yx}(a, b)h + f''_{y^2}(a, b)\theta_2 h + \\ + \varepsilon_2(h, \theta_2 h)\sqrt{1 + \theta_2^2}\,|h|,$$

$$f'_y(a, b + \theta_2 h) = f'_y(a, b) + f''_{y^2}(a, b)\theta_2 h + \varepsilon_2(0, \theta_2 h)\theta_2\,|h|.$$

En faisant la soustraction et en tenant compte de la formule (3), on obtient

$$\Delta(h) = f''_{yx}(a, b)h^2 + \beta(h)h^2$$

où $\beta(h) = \pm[\varepsilon_2(h, \theta_1 h)\sqrt{1 + \theta_2^2} - \varepsilon_2(0, \theta_2 h)\theta_2] \to 0$ lorsque $h \to 0$. Par conséquent,

$$(4) \qquad f''_{yx}(a, b) = \lim_{h \to 0} \frac{\Delta(h)}{h^2}.$$

En comparant (2) et (4), on conclut que $f''_{xy}(a, b) = f''_{yx}(a, b)$.

REMARQUE. Les théorèmes 1 et 2 restent en vigueur si f dépend, en plus de x et y, d'autres arguments. En effet, ces arguments restent constants dans les dérivations (par rapport à x et y) qu'on utilise dans les démonstrations de ces théorèmes.

4. Dérivées partielles d'ordre *n*

Nous avons défini les dérivées partielles d'ordres 1 et 2. Par récurrence, on définit les dérivées partielles de n'importe quel ordre supérieur (bien sûr, là où elles existent). Notamment, les dérivées partielles d'ordre n sont définies en tant que dérivées partielles d'ordre 1 des dérivées partielles d'ordre $n - 1$. Il est aussi rationnel d'entendre par dérivée partielle d'ordre nul de f la fonction f elle-même.

Suivant le théorème 2, la différentiabilité de f'_x et f'_y entraîne l'égalité des dérivées mixtes f''_{xy} et f''_{yx}, i.e. la commutativité des dérivations par rapport à x et y. Est-ce qu'on peut généraliser ce résultat aux dérivées partielles d'ordres supérieurs ? Par exemple, est-ce que (dans les conditions appropriées) $f^{\mathrm{IV}}_{xyxy} = f^{\mathrm{IV}}_{yyxx}$?

THÉORÈME 3. *Supposons que toutes les dérivées partielles d'ordre n de la fonction $f(x, y)$ existent et sont différentiables. Il s'ensuit que pour tout entier $k \in [2, n + 1]$ les dérivations par rapport à x et y dans chaque dérivée partielle d'ordre k sont permutables.*

Démonstration. Etant donné que toutes les dérivées partielles d'ordre n de la fonction f sont différentiables, il existe pour f toutes les dérivées partielles d'ordre $n + 1$ (corollaire du théorème 6.3.5). Il existe, bien sûr, toutes les dérivées partielles d'ordre $< n$. Elles sont différentiables. En effet, vu que les deux dérivées partielles d'ordre 1 de chaque dérivée partielle d'ordre $n - 1$ (qui sont les dérivées partielles d'ordre n de la fonction f) sont différentiables, elles sont continues (voir théorème 6.3.3) ; il en découle alors que les dérivées partielles d'ordre $n - 1$ sont différentiables (voir théorème 6.3.6). On en déduit de façon analogue la différentiabilité des dérivées partielles d'ordre $n - 2$, etc., jusqu'à f'_x et f'_y y compris. Soit maintenant $k \in [2, n + 1]$ et soit $f^{(k)}_{z_1 \ldots z_k}$, où tout z_i est x ou y, une dérivée partielle d'ordre k. Soit $z_i \neq z_{i+1}$, disons $z_i = y$ et $z_{i+1} = x$. Comme $f^{(i)}_{z_1 \ldots z_{i-1}x}$ et $f^{(i)}_{z_1 \ldots z_{i-1}y}$ sont différentiables, on a $f^{(i+1)}_{z_1 \ldots z_{i-1}xy} = f^{(i+1)}_{z_1 \ldots z_{i-1}yx}$ en vertu du théorème 2. Donc, $f^{(k)}_{z_1 \ldots z_{i-1}yxz_{i+2} \ldots z_k} = f^{(k)}_{z_1 \ldots z_{i-1}xyz_{i+2} \ldots z_k}$. Ainsi donc, deux dérivations quelconques voisines par rapport à x et y sont permutables. Or toute permutation de dérivations par rapport à x et y peut être effectuée à l'aide d'un nombre fini de permutations des dérivations voisines, par exemple,

$$f^{\mathrm{IV}}_{xyxy} = f^{\mathrm{IV}}_{yxxy} = f^{\mathrm{IV}}_{yxyx} = f^{\mathrm{IV}}_{yyxx}.$$

Il s'ensuit que toutes dérivations par rapport à x et y sont permutables.

COROLLAIRE. *Si les hypothèses du théorème* 3 *sont satisfaites, la fonction f peut avoir seulement $k + 1$ dérivées partielles distinctes d'ordre k.*

En effet, toute dérivée partielle d'ordre k peut être ramenée par permutation de dérivations à la dérivée partielle où sont d'abord faites toutes les dérivations par rapport à x et ensuite celles par rapport à y. On n'obtient que $k + 1$ dérivées partielles

$$f^{(k)}_{x^k},\ f^{(k)}_{x^{k-1}y},\ \ldots,\ f^{(k)}_{xy^{k-1}},\ f^{(k)}_{y^k}.$$

Le théorème analogue est aussi vrai pour les fonctions dont le nombre de variables est supérieur à deux.

QUESTION. Combien de dérivées partielles d'ordre k a une fonction $k - 1$ fois différentiable de trois variables ? de n variables ?

§ 2. Différentielles d'ordres supérieurs. Formule de Taylor

1. Différentielles d'ordres supérieurs

Soit $f(x, y)$ une fonction différentiable (par conséquent, D_f est un ensemble ouvert). Conformément au n° 6.3.2, on appelle *différentielle* de f au point $(x, y) \in D_f$ une fonction linéaire (ou comme on dit encore, *forme linéaire*)

$$\mathrm{d}f(x, y)(h, k) = \partial_1 f(x, y)h + \partial_2 f(x, y)k$$

de h et k (qui peuvent prendre n'importe quelles valeurs réelles indépendamment l'une de l'autre). En fixant h et k, on peut envisager d comme une opération $h\partial_1 + k\partial_2$ sur f qui consiste en dérivations ∂_1 et ∂_2, en multiplication des dérivées partielles obtenues par h et k respectivement et en addition de ces produits. Ceci étant, l'opération $\mathrm{d} = h\partial_1 + k\partial_2$ est linéaire, i.e. $\mathrm{d}(f + g) = \mathrm{d}f + \mathrm{d}g$ et $\mathrm{d}(\lambda f) = \lambda \mathrm{d}f$. Si $\partial_1 f$ et $\partial_2 f$ sont, à leur tour, différentiables, on peut répéter l'opération en formant $\mathrm{d}^2 f = \mathrm{d}(\mathrm{d}f)$:

$$\mathrm{d}^2 f = (h\partial_1 + k\partial_2)^2 f = (h\partial_1 + k\partial_2)(h\partial_1 f + k\partial_2 f) = \\ = h^2\partial_1^2 f + hk\partial_1\partial_2 f + kh\partial_2\partial_1 f + k^2\partial_2^2 f.$$

Comme, en vertu du théorème 2 du paragraphe précédent, les dérivations ∂_1 et ∂_2 sont permutables, on obtient (en notation habituelle)

$$\text{(1)} \quad \mathrm{d}^2 f(x, y)(h, k) = \frac{\partial^2 f}{\partial x^2}(x, y)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x, y)hk + \frac{\partial^2 f}{\partial y^2}(x, y)k^2.$$

Si $\partial_1 f$ et $\partial_2 f$ sont différentiables, la forme quadratique (1) de h et k est appelée *différentielle d'ordre* 2 de la fonction f au point (x, y).

Si les dérivées partielles d'ordre 2 de la fonction f sont différentiables, on peut aussi obtenir sa différentielle d'ordre 3. En général, si toutes les dérivées partielles d'ordre $n - 1$ de la fonction f existent et sont différentiables (et par suite, il existe toutes les dérivées partielles d'ordre n), on appelle *différentielle d'ordre* n la forme de degré n

$$\text{(2)} \qquad \mathrm{d}^n f(x, y)(h, k) = (h\partial_1 + k\partial_2)^n f(x, y).$$

Ceci étant,

$$\text{(3)} \qquad (h\partial_1 + k\partial_2)^n f(x, y) = \sum_{i=0}^{n} C_n^i \frac{\partial^n f}{\partial x^{n-i}\partial y^i}(x, y)h^{n-i}k^i,$$

où C_n^i sont les coefficients binomiaux. En effet, vu que les opérations ∂_1 et ∂_2 sont linéaires et permutables, on démontre, tout comme si ∂_1 et ∂_2 étaient des nombres, que

$$\mathrm{d}^n(h\partial_1 + k\partial_2)^n = \sum_{i=0}^{n} C_n^i \partial_1^{n-i}\partial_2^i h^{n-i}k^i.$$

Posons $\varrho = \sqrt{h^2 + k^2}$. Il existe un angle φ tel que $h = \varrho \cos\varphi$, $k = \varrho \sin\varphi$. C'est pourquoi, on peut mettre les formules (2) et (3) sous la forme suivante :

$$\text{(4)} \qquad \mathrm{d}^n f(x, y)(h, k) = \left(\sum_{i=0}^{n} C_n^i \frac{\partial^n f}{\partial x^{n-i}\partial y^i}(x, y)\cos^{n-i}\varphi \sin^i\varphi\right)\varrho^n.$$

La fonction de φ située dans le second membre de la formule (4) étant bornée pour tout point fixe (x, y), cette formule montre que l'ordre de petitesse par rapport à ϱ de $\mathrm{d}^n f(x, y)(h, k)$ est au moins égal à n.

Notons que la notion de différentielle d'ordre supérieur se généralise aux fonctions de n'importe quel nombre de variables. Donc,

$$d^n f(x, y, z)(h, k, l) = (h\partial_1 + k\partial_2 + l\partial_3)^n f(x, y, z) \quad (h, k, l \in \mathbf{R}).$$

2. Formule de Taylor

Pour la fonction d'une variable, la formule de Taylor avec le reste de Lagrange donne un développement approximatif de $f(x + h)$ suivant les puissances de h :

$$(5) \quad f(x+h) = f(x) + f'(x)h + \frac{1}{2} f''(x)h^2 + \dots$$
$$\dots + \frac{1}{n!} f^{(n)}(x)h^n + \frac{1}{(n+1)!} f^{(n+1)}(x+\theta h)h^{n+1},$$

où θ est un nombre compris entre 0 et 1. Comme $f^{(k)}(x)h^k = d^k f(x)(h)$, la formule (5) peut être écrite comme suit :

$$f(x+h) = f(x) + df(x)(h) + \frac{1}{2!} d^2 f(x)(h) + \dots$$
$$\dots + \frac{1}{n!} d^n f(x)(h) + \frac{1}{(n+1)!} d^{n+1} f(x+\theta h)(h).$$

Une formule analogue est vraie pour les fonctions de tout nombre de variables. Déduisons-la pour le cas de deux variables.

THÉORÈME 1. *Supposons que toutes les dérivées partielles d'ordre n ($n = 0, 1, 2, \dots$) de la fonction $f : \mathbf{R}^2 \to \mathbf{R}$ existent et soient différentiables dans un voisinage U du point (x, y). Alors pour tous les h et k suffisamment petits,*

$$(6) \quad f(x+h, y+k) = f(x, y) + df(x, y)(h, k) +$$
$$+ \frac{1}{2!} d^2 f(x, y)(h, k) + \dots + \frac{1}{n!} d^n f(x, y)(h, k) +$$
$$+ \frac{1}{(n+1)!} d^{n+1} f(x+\theta h, y+\theta k)(h, k)$$

où θ est un nombre situé entre 0 et 1 pour chaque couple (h, k).

Démonstration. Le voisinage U contient un voisinage circulaire suffisamment petit $U((x, y) ; r)$ du point (x, y). Si $h^2 + k^2 < r^2$, les points $(x + th, y + tk)$ appartiennent à $U((x, y) ; r)$, et par conséquent, à U pour tout $t \in [0, 1]$, de sorte que f est définie en ces points et y possède toutes les dérivées partielles d'ordre n qui sont différentiables. Ayant fixé h et k, introduisons une fonction auxiliaire

$$(7) \quad F(t) = f(x + th, y + tk) \quad (t \in [0, 1]).$$

Montrons qu'elle possède les dérivées jusqu'à l'ordre $n + 1$ et que

$$F^{(m)}(t) = \mathrm{d}^m f(x + th, y + tk)(h, k) \tag{8}$$

pour tout entier $m \in [1, n + 1]$ et tout $t \in [0, 1]$.

En effet, $F(t) = f(X, Y)$ où $X = x + th$, $Y = y + tk$. Comme f, X et Y sont différentiables, F est différentiable et d'après la formule 7.2(12)

$$F'(t) = \partial_1 f(X, Y)\frac{\mathrm{d}X}{\mathrm{d}t} + \partial_2 f(X, Y)\frac{\mathrm{d}Y}{\mathrm{d}t} = \partial_1 f(X, Y)h + \\ + \partial_2 f(X, Y)k = (h\partial_1 + k\partial_2)f(X, Y) = \mathrm{d}f(X, Y)(h, k).$$

Supposons qu'il soit établi que F possède la dérivée m-ième pour un $m < n + 1$ et que

$$F^{(m)}(t) = \mathrm{d}^m f(X, Y)(h, k) = \sum_{i=0}^{m} C_m^i \frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)^{m-i}(h, k). \tag{8'}$$

Vu que f possède dans le disque $U((x, y) ; r)$ des dérivées partielles d'ordre n qui sont toutes différentiables, la démonstration du théorème 3 du paragraphe précédent montre que ceci est vrai pour toutes les dérivées partielles d'ordres inférieurs. C'est pourquoi, toutes les dérivées $\frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)$ sont des fonctions dérivables de t et, d'après la formule 7.2(12),

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y) = \partial_1 \frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)\frac{\mathrm{d}X}{\mathrm{d}t} + \\ + \partial_2 \frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)\frac{\mathrm{d}Y}{\mathrm{d}t} = (h\partial_1 + k\partial_2)\frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y).$$

Il résulte alors de la formule (8′) que $F^{(m)}(t)$ est différentiable et

$$F^{(m+1)}(t) = \sum_{i=0}^{m} C_m^i (h\partial_1 + k\partial_2)\frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)h^{m-i}k^i = \\ = (h\partial_1 + k\partial_2)\sum_{i=0}^{m} C_m^i \frac{\partial^m f}{\partial x^{m-i}\partial y^i}(X, Y)h^{m-i}k^i = \\ = (h\partial_1 + k\partial_2)\mathrm{d}^m f(X, Y)(h, k) = \mathrm{d}^{m+1} f(X, Y)(h, k).$$

Donc, F possède sur le segment $[0, 1]$ des dérivées jusqu'à l'ordre $n + 1$. Par conséquent, d'après la formule de Taylor pour une fonction d'une variable,

$$F(1) = F(0) + F'(0) + \frac{1}{2!}F''(0) + \ldots + \frac{1}{n!}F^{(n)}(0) + \frac{1}{(n+1)!}F^{(n+1)}(\theta)$$

où θ est un nombre compris entre 0 et 1. Mais en vertu de (7) et (8), c'est justement la formule (6).

Pour $n = 0$, la formule (6) prend la forme suivante

$$(9)\quad f(x+h, y+k) = \\ = f(x, y) + \left(\frac{\partial f}{\partial x}(x+\theta h, y+\theta k)h + \frac{\partial f}{\partial y}(x+\theta h, y+\theta k)k\right).$$

§ 3. Théorème de Ferma. Extrémums

1. Théorème de Ferma

THÉORÈME 1 (théorème de Ferma). *Si la fonction* $f: \mathbf{R}^n \to \mathbf{R}$ *prend au point intérieur* $x^{(0)}$ *de l'ensemble* $E \subset D_f$ *une valeur maximale ou minimale sur* E *et si elle est différentiable en ce point,* $f'(x^{(0)}) = 0$, *i.e. toutes les dérivées partielles d'ordre* 1 *de la fonction* f *au point* $x^{(0)}$ *s'annulent.*

Démonstration. Soit, par exemple, $x^{(0)}$ un point où f prend sa valeur maximale. Il existe alors un $r > 0$ tel que $U(x^{(0)}; r) \subset E$ et $f(x^{(0)}) \geqslant \geqslant f(x)$ pour tout $x \in U(x^{(0)}; r)$. Considérons la i-ième fonction partielle

$$\varphi_i(z) = f(x_1^{(0)}, \ldots, x_{i-1}^{(0)}, z, x_{i+1}^{(0)}, \ldots, x_n^{(0)}).$$

Conformément aux résultats du n° 6.1.2, elle est différentiable au point $x_i^{(0)}$ et $\varphi_i'(x_i^{(0)}) = (\partial_i f)(x^{(0)})$. On a en outre $\varphi_i(x^{(0)}) \geqslant \varphi_i(z)$ pour tout $z \in \in]x_i^{(0)} - r, x_i^{(0)} + r[$, de sorte que $\varphi_i'(x_i^{(0)}) = 0$. Ainsi donc, $(\partial_i f)(x^{(0)}) = = 0$ pour tout entier $i \in]1, n[$, et par suite, $f'(x^{(0)}) = 0$.

DÉFINITION 1. Le point x où la fonction $f: \mathbf{R}^n \to \mathbf{R}$ est différentiable et $f'(x) = 0$ est appelé *point stationnaire* de la fonction f.

Donc, *pour qu'au point intérieur* x *de l'ensemble* $E \subset D_f$ *la fonction* f *prenne sa valeur maximale ou minimale sur* E, *il est nécessaire* (*si* f *est différentiable en* x) *que* x *soit un point stationnaire de cette fonction.* Cependant, cette condition n'est pas en général suffisante. Ainsi, (0, 0) est un point stationnaire de la fonction $f(x, y) = xy$, mais il n'est ni le point de la valeur maximale ni le point de la valeur minimale de cette fonction dans aucun de ses voisinages. En effet, tout voisinage du point (0, 0) contient des points (x, y) pour lesquels $f(x, y) > f(0, 0)$ $(=0)$ (ce sont les points intérieurs au premier et au troisième quadrant) et des points pour lesquels $f(x, y) < f(0, 0)$ (ce sont les points intérieurs au deuxième et au quatrième quadrant) (fig. 22).

Cependant, le calcul de la valeur maximale ou minimale de la fonction différentiable f sur l'ensemble $E \subset D_f$ se ramène à la recherche d'un point stationnaire si l'on sait qu'une telle valeur existe et qu'elle est prise en un point intérieur de l'ensemble E ; ce point doit alors être stationnaire en vertu du théorème 1. La situation devient encore plus simple si E contient

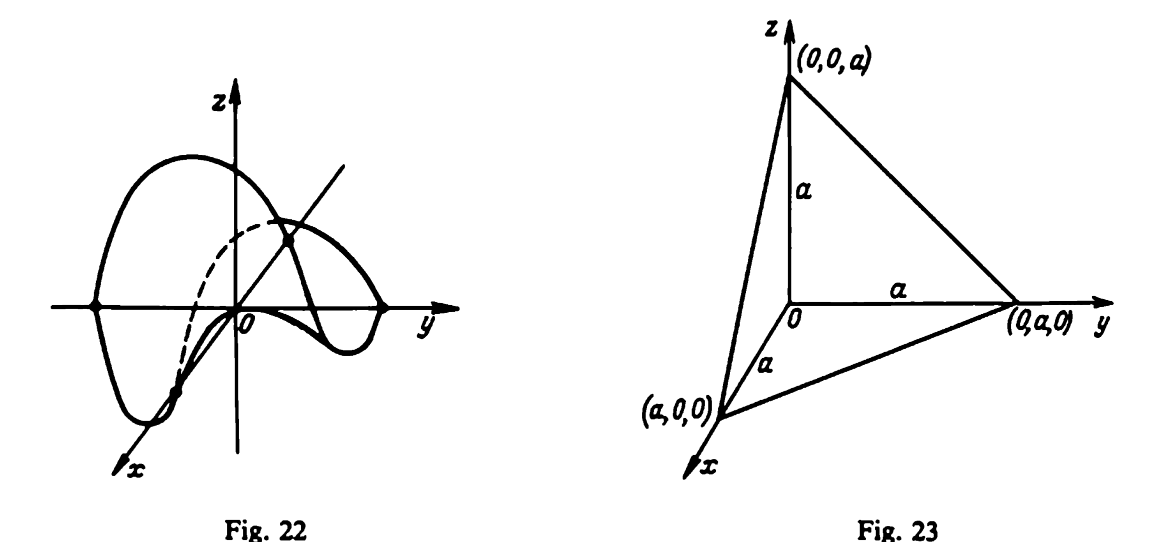


Fig. 22 Fig. 23

un seul point stationnaire ; ce point doit donc être le point recherché où la fonction présente sa valeur maximale ou minimale.

EXEMPLE 1. Il faut décomposer le nombre strictement positif donné a en somme de quatre nombres strictement positifs de telle sorte que leur produit soit maximal. En décomposant a en somme des variables x, y, z et $t = a - x - y - z$, on réduit le problème à la recherche de la valeur maximale de la fonction

$$f(x, y, z) = xyz(a - x - y - z)$$

sur un ensemble E défini par les inégalités

$$x > 0,\ y > 0,\ z > 0,\ x + y + z < a,$$

E étant l'intérieur de l'ensemble

$$T = \{(x, y, z) \in \mathbf{R}^3 \mid x \geqslant 0,\ y \geqslant 0,\ z \geqslant 0,\ x + y + z \leqslant a\},$$

i.e. d'un tétraèdre formé par les plans de coordonnées $x = 0, y = 0, z = 0$ et par le plan $x + y + z = a$ (fig. 23). En effet, il est évident que $E \subset T$. Si $(x_0, y_0, z_0) \in E$, on a $\eta = \min\{x_0, y_0, z_0, a - x_0 - y_0 - z_0\} > 0$ et le cube $Q((x_0, y_0, z_0) ; \eta)$ est contenu dans E, i.e. (x_0, y_0, z_0) est un point intérieur à E et d'autant plus à T, de sorte que $E \subset]T[$. Enfin, si $(x_0, y_0, z_0) \in T \setminus I$, de façon qu'au moins une des égalités $x_0 = 0, y_0 = 0, z_0 = 0$ ou $a - x_0 - y_0 - z_0 = 0$ est satisfaite, il existe dans tout voisinage du point (x_0, y_0, z_0) un point (x, y, z) satisfaisant à l'une des inégalités $x < 0, y < 0, z < 0$ ou $a - x - y - z < 0$ au moins et, par conséquent, n'appartenant pas à T, de sorte que $(x_0, y_0, z_0) \notin]T[$. Etant une intersection des demi-espaces fermés $x \geqslant 0, y \geqslant 0, z \geqslant 0$ et $x + y + z \leqslant a$, le tétraèdre T est fermé ; en outre, puisque les coordonnées x, y, z de ses points

sont contenues dans le segment $[0, a]$, il est borné ; par conséquent, T est un compact en vertu du théorème 3.2.13. Comme f est continue, elle possède sur T la valeur maximale (voir théorème 3.2.9). Celle-ci est prise dans E car $f(T \setminus E) = 0$ et $f(E) > 0$. Ainsi donc, la valeur maximale de la fonction f sur T est la valeur maximale recherchée de f sur l'intérieur E de T, et le point où cette valeur se présente est un point stationnaire de la fonction f. Or

$$f'_x(x, y, z) = yz(a - x - y - z) - xyz = yz(t - x),$$
$$f'_y(x, y, z) = xz(t - y), \qquad f'_z(x, y, z) = xy(t - z).$$

Par conséquent, le point cherché (x, y, z) de la valeur maximale satisfait aux équations

$$yz(t - x) = 0, \qquad xz(t - y) = 0, \qquad xy(t - z) = 0,$$

ou, puisque $xyz \neq 0$ en ce point, aux équations

$$t - x = 0, \qquad t - y = 0, \qquad t - z = 0.$$

Ceci étant, $x + y + z + t = a$, de sorte que $x = y = z = t = a/4$.

Nous avons encore montré en passant que quels que soient les nombres strictement positifs x, y, z, t,

$$xyzt \leqslant \left(\frac{x + y + z + t}{4}\right)^4, \text{ i.e. } \sqrt[4]{xyzt} \leqslant \frac{x + y + z + t}{4}$$

(poser $x + y + z + t = a$), l'égalité ayant lieu si $x = y = z = t$.

EXERCICE 1. Décomposer le nombre strictement positif a en somme de n termes strictement positifs de telle sorte que leur produit soit maximal.

EXEMPLE 2. Il faut représenter le nombre strictement positif donné a sous forme d'un produit de quatre nombres strictement positifs de telle sorte que leur somme soit minimale. En représentant a sous forme du produit des facteurs variables x, y, z et $\dfrac{a}{xyz}$, on réduit le problème à la recherche de la valeur minimale de la fonction

$$f(x, y, z) = x + y + z + \frac{a}{xyz}$$

sur l'ensemble $E = \{(x, y, z) \in \mathbf{R}^3 \mid x > 0,\ y > 0,\ z > 0\}$. Soit c l'une des valeurs prises par la fonction f sur E et soit $K = \{(x, y, z) \in E \mid f(x, y, z) \leqslant c\}$. Si $(x, y, z) \in K$, on a $x + y + z < f(x, y, z) \leqslant c$ et comme $x > 0, y > 0, z > 0$, on obtient $x < c, y < c, z < c$. Ainsi donc, K est borné. D'autre part, comme $\dfrac{a}{xyz} < f(x, y, z) \leqslant c$, on a (puisque

$y < c$ et $z < c$) $x > \dfrac{a}{cyz} > \dfrac{a}{c^3}$ et aussi $y > \dfrac{a}{c^3}$ et $z \geqslant \dfrac{a}{c^3}$. Pour cette raison, si $(x_n, y_n, z_n) \in K$ et $(x_n, y_n, z_n) \to (x, y, z)$, on peut écrire $x, y, z \geqslant \dfrac{a}{c^3}$, de sorte que $(x, y, z) \in E$ et alors $f(x, y, z) =$ $= \lim f(x_n, y_n, z_n) \leqslant c$, donc $(x, y, z) \in K$. Ainsi, K est fermé. Par suite, K est un compact et f présente sur K la valeur minimale prise en un point (x_0, y_0, z_0). Mais alors $f(x_0, y_0, z_0)$ est la valeur minimale de f sur tout l'ensemble E. En effet, si $(x, y, z) \in K$, on a $f(x_0, y_0, z_0) \leqslant f(x, y, z)$ d'après la définition du point (x_0, y_0, z_0), et si $(x, y, z) \in E \setminus K$, on obtient $f(x, y, z) > c \geqslant f(x_0, y_0, z_0)$. Comme E est un ouvert et la fonction f est différentiable, (x_0, y_0, z_0) est un point stationnaire de f (voir théorème 1), i.e.

$$f'_x(x_0, y_0, z_0) = 1 - \frac{a}{x_0^2 y_0 z_0} = 0, \qquad f'_y(x_0, y_0, z_0) = 1 - \frac{a}{x_0 y_0^2 z_0} = 0,$$

$$f'_z(x_0, y_0, z_0) = 1 - \frac{a}{x_0 y_0 z_0^2} = 0.$$

D'où il résulte aisément que $x_0 = y_0 = z_0 = \dfrac{a}{x_0 y_0 z_0} = \sqrt[4]{a}$.

On a de nouveau montré que quels que soient les nombres strictement positifs x, y, z, t,

$$x + y + z + t \geqslant 4\sqrt[4]{xyzt}, \text{ i.e. } \sqrt[4]{xyzt} \leqslant \frac{x + y + z + t}{4}$$

(poser $xyzt = a$), avec la même remarque concernant l'égalité.

2. Points d'extrémum

DÉFINITION 2. Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$. On dit que le point $x^{(0)}$ est un *point de maximum* (resp. *de minimum*) et que $f(x^{(0)})$ est le *maximum* (resp. le *minimum*) si $x^{(0)}$ possède un voisinage $U \subset D_f$ tel que $f(x^{(0)}) \geqslant f(x)$ (resp. $\leqslant f(x)$) pour tous les points $x \in U$. Si par ailleurs $f(x^{(0)}) > f(x)$ (resp. $< f(x)$) pour tout $x \in U$ distinct de $x^{(0)}$, on dit que $x^{(0)}$ est un *point de maximum strict* (resp. *de minimum strict*). Les points de minimum et les points de maximum de la fonction f sont appelés *points d'extrémum*.

Il résulte directement du théorème 1 le fait suivant : *pour que le point x où la fonction f est différentiable soit un point d'extrémum de cette fonction, il est nécessaire que x soit un point stationnaire, i.e. que $f'(x) = 0$.* On a vu également que cette condition n'est pas suffisante. Ainsi donc, il faut avoir des conditions permettant de savoir quand le point stationnaire donné est un point d'extrémum et de quel extrémum il s'agit : de maximum

ou de minimum. Nous n'analyserons que le cas d'une fonction de deux variables.

THÉORÈME 2. *Supposons que la fonction* $f: \mathbf{R}^2 \to \mathbf{R}$ *soit différentiable dans un voisinage de son point stationnaire* (x_0, y_0) *et que* f'_x *et* f'_y *soient différentiables en ce point (de sorte que suivant le théorème 2 du § 1,* f''_{xy} *et* f''_{yx} *y coïncident). Posons*

$$A = f''_{x^2}(x_0, y_0), \qquad B = f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0), \qquad C = f''_{y^2}(x_0, y_0)$$

et soit $AC - B^2 \neq 0$. *Alors :*

1) *si* $AC - B^2 > 0$, (x_0, y_0) *est un point de maximum strict de f pour* $A < 0$ *et de minimum strict pour* $A > 0$;

2) *si* $AC - B^2 < 0$, (x_0, y_0) *n'est ni point de maximum ni point de minimum.*

Démonstration. La fonction f est différentiable dans un r-voisinage du point (x_0, y_0). Soit V un ensemble de tous les couples $(h, k) \in \mathbf{R}^2$ satisfaisant à la double inégalité $0 < \varrho < r$ où $\varrho = \sqrt{h^2 + k^2}$. D'après la formule (9) du § 2 (qu'on peut appliquer parce que f satisfait aux hypothèses du théorème 1 du § 2 pour $n = 0$) il existe pour chaque couple $(h, k) \in V$ un $\theta \in]0, 1[$ tel que

$$(1) \qquad \Delta f = f(x_0 + h, y_0 + k) - f(x_0, y_0) = \\ = f'_x(x_0 + \theta h, y_0 + \theta k)h + f'_y(x_0 + \theta h, y_0 + \theta k)k.$$

Si pour un (h, k) il existe plusieurs θ, on choisit l'un d'eux ; ainsi donc, θ est une fonction de (h, k) sur V. Comme f'_x est différentiable au point (x_0, y_0) et s'y annule (puisque (x_0, y_0) est un point stationnaire), on a

$$f'_x(x, y) = A(x - x_0) + B(y - y_0) + \\ + \alpha_1(x - x_0, y - y_0)\sqrt{(x - x_0)^2 + (y - y_0)^2}$$

où $\alpha_1 \to 0$ quand $(x - x_0, y - y_0) \to (0, 0)$. En tenant compte de ce que $(\theta h, \theta k) \to 0$ lorsque $(h, k) \to (0, 0)$, on obtient en particulier que

$$(2) \qquad f'_x(x_0 + \theta h, y_0 + \theta k) = \theta(Ah + Bk + \alpha_1\varrho)$$

où $\alpha_1 = \alpha_1(\theta h, \theta k) \to 0$ lorsque $(h, k) \to 0$. De façon analogue,

$$(3) \qquad f'_y(x_0 + \theta h, y_0 + \theta k) = \theta(Bh + Ck + \alpha_2\varrho)$$

où $\alpha_2 = \alpha_2(\theta h, \theta k) \to 0$ lorsque $(h, k) \to (0, 0)$. En substituant (2) et (3) dans (1), on obtient

$$(4) \qquad \Delta f = \theta[(Ah^2 + 2Bhk + Ck^2) + (\alpha_1 h + \alpha_2 k)\varrho].$$

Pour tout couple (h, k) il existe un $\varphi \in [-\pi, \pi]$ tel que

$$h = \varrho\cos\varphi, \qquad k = \varrho\sin\varphi.$$

En posant

$$F(\varphi) = A\cos^2\varphi + 2B\cos\varphi\sin\varphi + C\sin^2\varphi,$$

on donne à la formule (4) la forme suivante

$$\Delta f = \theta\varrho^2[F(\varphi) + \alpha] \tag{5}$$

où $\alpha = \alpha_1\cos\varphi + \alpha_2\sin\varphi$, de sorte que $|\alpha| \leqslant |\alpha_1| + |\alpha_2|$. C'est pourquoi,

$$(\forall\varepsilon > 0)\ (\exists\delta \in]0, r])\ (\varrho < \delta \Rightarrow |\alpha| < \varepsilon). \tag{6}$$

Montrons que si $F(\varphi) > 0$ (resp. <0) pour tout φ, (x_0, y_0) est un point de minimum strict (resp. de maximum strict), et si $F(\varphi)$ prend des valeurs négatives et positives, (x_0, y_0) n'est ni point de minimum ni point de maximum. En effet, étant une fonction continue sur $[-\pi, \pi]$, F présente sur cet intervalle la valeur minimale c et la valeur maximale d. Si $F(\varphi) > 0$ pour tout φ, on a, en particulier, $c > 0$. Par suite, en vertu de (6),

$$(\exists\delta \in]0, r])\ (\varrho < \delta \Rightarrow |\alpha| < c \Rightarrow \alpha > -c).$$

Comme $F(\varphi) > c$, on obtient $F(\varphi) + \alpha > 0$ et il résulte de la formule (5) que $\Delta f > 0$ pour tout $(h, k) \in V$ tel que $\varrho < \delta$, de sorte que (x_0, y_0) est un point de maximum strict. De façon analogue, si $F(\varphi) < 0$ pour tout φ, on a, en particulier, $d < 0$. C'est pourquoi,

$$(\exists\delta \in]0, r])\ (\varrho < \delta \Rightarrow |\alpha| < -d \Rightarrow \alpha < -d).$$

Puisque $F(\varphi) \leqslant d$, on obtient $F(\varphi) + \alpha < 0$ et il résulte de la formule (5) que $\Delta f < 0$ pour tout $(h, k) \in V$ tel que $\varrho < \delta$, de sorte que (x_0, y_0) est un point de maximum strict. Enfin, si F prend des valeurs de signes différents, i.e. s'il existe des $\varphi_1, \varphi_2 \in [-\pi, \pi]$ tels que $F(\varphi_1) < 0 < F(\varphi_2)$, on peut écrire

$$(\exists\delta \in]0, r])\ (\varrho < \delta \Rightarrow |\alpha| < \min\{|F(\varphi_1)|, F(\varphi_2)\}).$$

Dans ce cas, $F(\varphi_1) + \alpha < 0 < F(\varphi_2) + \alpha$ et il résulte de la formule (5) que $f(x_0 + \varrho\cos\varphi_1,\ y_0 + \varrho\sin\varphi_1) < f(x_0,\ y_0) < f(x_0 + \varrho\cos\varphi_2,\ y_0 + \varrho\sin\varphi_2)$ pour tout $\varrho \in]0, \delta[$, de sorte que (x_0, y_0) n'est ni point de minimum ni point de maximum puisque ϱ est aussi petit que l'on veut. Ainsi donc, le problème se réduit à l'étude du signe de $F(\varphi)$.

Soit $A \neq 0$. Alors

$$F(\varphi) = \frac{1}{A}[(A\cos\varphi + B\sin\varphi)^2 + (AC - B^2)\sin^2\varphi]. \tag{7}$$

Les termes de la somme entre les crochets ne peuvent pas s'annuler simultanément : si $(AC - B^2)\sin^2\varphi = 0$, on a $\sin\varphi = 0$ (car $AC - B^2 \neq 0$ par

hypothèse), de sorte que $\cos\varphi = \pm 1$ et, par suite, $(A\cos\varphi + B\sin\varphi)^2 = A^2 \neq 0$. Pour cette raison, $AF(\varphi) > 0$ pour tout φ si $AC - B^2 > 0$, i.e. $F(\varphi)$ et A sont de même signe, et donc, en vertu de ce qui était dit plus haut, (x_0, y_0) est un point de maximum strict quand $A < 0$ et de minimum strict quand $A > 0$. Si $AC - B^2 < 0$, les valeurs de $F(\varphi)$ sont de signes opposés pour les valeurs données de φ puisque

$$(A\cos\varphi + B\sin\varphi)^2 + (AC - B^2)\sin^2\varphi = \begin{cases} (AC - B^2)\sin^2\varphi < 0 & \text{pour} \quad \varphi = \operatorname{Arctg}\left(-\dfrac{B}{A}\right), \\ A^2 > 0 & \text{pour} \quad \varphi = 0, \end{cases}$$

de sorte que (x_0, y_0) n'est ni point de minimum ni point de maximum.

Soit maintenant $A = 0$. Comme $AC - B^2 \neq 0$, on obtient $B \neq 0$, de sorte que $AC - B^2 = -B^2 < 0$. Dans ce cas,

$$F(\varphi) = 2B\cos\varphi\sin\varphi + C\sin^2\varphi = (2B\cos\varphi + C\sin\varphi)\sin\varphi.$$

Mais l'expression entre parenthèses tend vers $2B$ quand $\varphi \to 0$ et, pour cette raison, a le signe de B pour les valeurs de φ suffisamment petites en valeur absolue. D'autre part, $\sin\varphi$ change de signe en passant par 0. Par conséquent, F a des valeurs positives et négatives, si bien que (x_0, y_0) n'est, cette fois encore, un point d'extrémum.

REMARQUE. Le théorème 2 laisse de côté le cas où $AC - B^2 = 0$ (si les hypothèses relatives à f, f'_x et f'_y sont satisfaites). Si $A \neq 0$, on obtient, en vertu de la formule (7),

$$F(\varphi) = \frac{1}{A}(A\cos\varphi + B\sin\varphi)^2,$$

de sorte que $F(\varphi) > 0$ pour tout φ, excepté $\varphi = \operatorname{Arctg}\left(-\dfrac{B}{A}\right)$ où $F(\varphi) = 0$. Comme on le voit de la formule (5), le signe de Δf est défini pour ce φ par le signe de α ; pour définir ce dernier, il faut des données supplémentaires. Si $A = 0$, de sorte que $B = 0$ (puisque $AC - B^2 = 0$), on voit que $F(\varphi) = C\sin^2\varphi$ et la situation est analogue : le signe de Δf est défini pour $\varphi = 0$ par un signe de α que nous ignorons.

EXEMPLE 3. Etudions les extrémums de la fonction

$$f(x, y) = x^2 - y^2 + e^{1-x^2}.$$

En résolvant les équations

$$f'_x(x, y) = 2x(1 - e^{1-x^2}) = 0, \qquad f'_y(x, y) = -2y = 0,$$

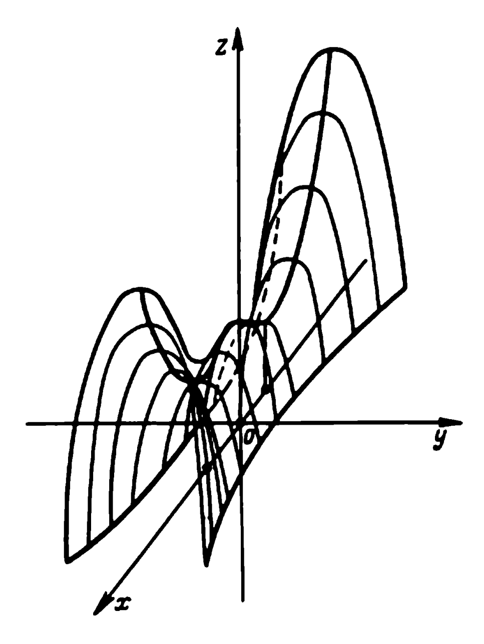


Fig. 24

on trouve trois points stationnaires : (0, 0), (1, 0), (−1, 0). Vu que

$$f''_{x^2}(x, y) = 2 - 2e^{1-x^2} + 4x^2e^{1-x^2}, \qquad f''_{xy}(x, y) = 0,$$
$$f''_{y^2}(x, y) = -2,$$

$AC - B^2$ a en ces points les valeurs respectives $-4(1 - e)$, -8, -8. Par conséquent, d'après le théorème 2, (1, 0) et (−1, 0) ne sont pas des points d'extrémum, et (0, 0) est un point de maximum strict (puisque $f''_{x^2}(0, 0) = 2 - 2e < 0$). Ainsi donc, la fonction étudiée, définie sur tout le $\mathbf{R}^2$, a un seul point d'extrémum, savoir le point de maximum strict (0, 0). Or $e = f(0, 0)$ n'est pas la valeur maximale de la fonction : $f(x, 0) = x^2e^{1-x^2}$ prend des valeurs aussi grandes que l'on veut. (Il est à noter que si une fonction d'une seule variable, définie et dérivable sur un intervalle, possède sur cet intervalle un seul point d'extrémum, ce point est aussi le point de la valeur maximale ou minimale de la fonction selon qu'il est le point de maximum ou de minimum respectivement.) Le graphe de la fonction étudiée est représenté à la fig. 24.

EXEMPLE 4. Etudions les extrémums de la fonction

$$f(x, y) = x^4 + y^4 - (x + y)^2 = x^4 + y^4 - x^2 - 2xy - y^2.$$

En résolvant les équations

$$f'_x(x, y) = 4x^3 - 2x - 2y = 0,$$
$$f'_y(x, y) = 4y^3 - 2x - 2y = 0,$$

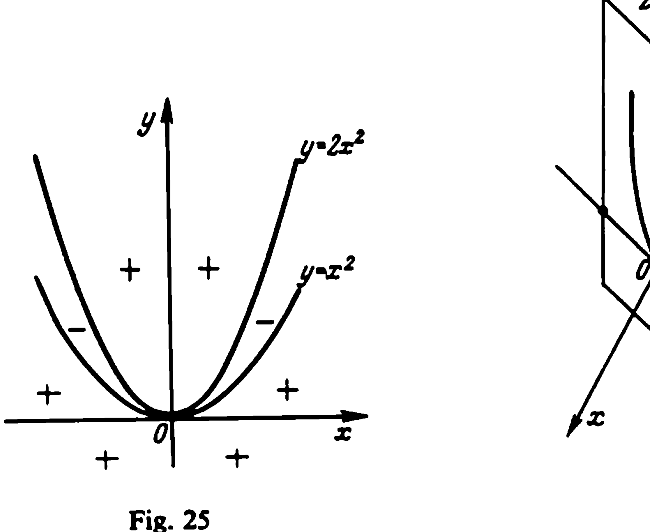


Fig. 25

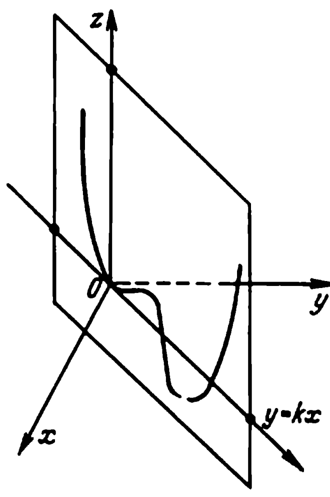


Fig. 26

on trouve trois points stationnaires : (0, 0), (1, 1), $(-1, -1)$. Vu que

$$f''_{x^2}(x, y) = 12x^2 - 2, \qquad f''_{xy}(x, y) = -2, \qquad f''_{y^2}(x, y) = 12y^2 - 2,$$

$AC - B^2$ a en ces points les valeurs respectives 0, 96, 96. Par suite, d'après le théorème 2 (puisque $f''_{x^2}(\pm 1, \pm 1) > 0$), (1, 1) et $(-1, -1)$ sont des points de minimum strict. Le caractère du point stationnaire (0, 0) n'est pas élucidé dans le théorème 2 vu que $AC - B^2 = 0$. Or $f(x, -x) = 2x^4 > 0$ pour tout $x \neq 0$, et $f(x, x) = 2x^4 - 4x^2 < 0$ si $0 < |x| < 2$. Par conséquent, (0, 0) n'est pas le point d'extrémum. Ainsi donc, la fonction étudiée, définie sur le $\mathbf{R}^2$ entier, a justement deux points d'extrémum, tous les deux étant des points de minimum strict. (La fonction d'une seule variable, définie et dérivable sur un intervalle, a au moins un point de maximum situé entre deux points de minimum quelconques.)

EXEMPLE 5. La fonction

$$f(x, y) = (x - y)^2 + y^n, \tag{8}$$

où $n > 2$ est un entier, a un seul point stationnaire (0, 0). On a $AC - B^2 = 0$ en ce point, de sorte que son caractère n'est pas élucidé par le théorème 2. Mais on voit de la formule (8) que si n est un nombre pair, $f(x, y) > f(0, 0) = 0$ pour tout $(x, y) \neq (0, 0)$, si bien que (0, 0) est un point de minimum strict ; si par contre, n est un nombre impair, $f(x, x) > 0$ pour $x > 0$ et $f(x, x) < 0$ pour $x < 0$, de sorte que (0, 0) n'est pas un point d'extrémum.

EXEMPLE 6. La fonction

$$f(x, y) = (x^2 - y)(2x^2 - y)$$

s'annule sur les paraboles $y = x^2$ et $y = 2x^2$, elle est strictement positive au-dessous de la première parabole et au-dessus de la deuxième et strictement négative entre elles (fig. 25). Ainsi donc, dans tout voisinage du point (0, 0) cette fonction a des valeurs positives et négatives, et donc (0, 0) n'est pas son point d'extrémum. Cependant, la restriction de f à chaque droite passant par le point (0, 0) est strictement positive dans un voisinage pointé de ce point ; c'est pourquoi, (0, 0) est un point de minimum de f. Le graphe de la restriction de f à la droite $y = kx$ est représenté à la fig. 26 ; plus k est petit, plus la partie du graphe où $f < 0$ est proche du point (0, 0) et moins elle est grande.

CHAPITRE 9

APPLICATIONS DIFFÉRENTIABLES LOCALEMENT INVERSIBLES

§ 1. Applications linéaires inversibles

1. Notion d'application linéaire inversible

Soient X et Y des espaces vectoriels sur un même corps commutatif K. L'application linéaire $A : X \to Y$ est dite *inversible* s'il existe une application linéaire $B : Y \to X$ telle que $BA = I_X$ et $AB = I_Y$ où I_X et I_Y sont les opérateurs unitaires des espaces X et Y respectivement. On dit alors que B est l'*inverse* de A et on le note A^{-1}.

LEMME 1. *Une application linéaire $A : X \to Y$ est inversible si et seulement si elle est bijective, i.e. si elle est une injection de X sur Y.*

Démonstration. Supposons que A soit inversible et que B soit une application linéaire de Y dans X, inverse à A. Si $Ax_1 = Ax_2$, on a $x_1 = BAx_1 = BAx_2 = x_2$, si bien que A est injective. En outre, tout $y \in Y$ peut être présenté sous la forme $A(By)$ et appartient donc à $A(X)$, de sorte que $A(X) = Y$. Ainsi donc, A est bijective. Inversement : supposons que A soit bijective. Il existe alors pour tout $y \in Y$ un $x \in X$ et un seul tel que $y = Ax$. Mettons $x = By$. On démontre que B est une application de Y dans X, inverse à A, i.e. $BA = I_X$, $AB = I_Y$. Il ne reste qu'à démontrer que B est linéaire. Soient $y_1, y_2 \in Y$. Posons $x_1 = By_1$, $x_2 = By_2$, de sorte que $y_1 = Ax_1$, $y_2 = Ax_2$. Comme A est linéaire,

$$\lambda_1 y_1 + \lambda_2 y_2 = A(\lambda_1 x_1 + \lambda_2 x_2)$$

pour tous $\lambda_1, \lambda_2 \in K$. Mais alors

$$B(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 x_1 + \lambda_2 x_2 = \lambda_1 By_1 + \lambda_2 By_2,$$

i.e. B est linéaire et par là même A est inversible en tant qu'application linéaire.

LEMME 2. *Une application linéaire A est injective si et seulement si $Ax = 0 \Rightarrow x = 0$.*

Démonstration. Supposons que A soit injective et $Ax = 0$. Comme $A0 = 0$, on obtient $x = 0$. Inversement : soit $Ax = 0 \Rightarrow x = 0$. Si $Ax_1 = Ax_2$, i.e. $A(x_1 - x_2) = 0$, on a $x_1 - x_2 = 0$, i.e. $x_1 = x_2$, donc A est injective.

LEMME 3. *Toute application linéaire injective $A : X \to Y$ envoie les systèmes de vecteurs linéairement indépendants en systèmes de vecteurs linéairement indépendants.*

Démonstration. Soient $x_1, \ldots, x_k$ des vecteurs linéairement indépendants dans X et soit $\sum_{i=1}^{k} \lambda_i A x_i = 0$. Comme A est linéaire, $\sum_{i=1}^{k} \lambda_i A x_i = A\left(\sum_{i=1}^{k} \lambda_i x_i\right)$ et, par conséquent, $A\left(\sum_{i=1}^{k} \lambda_i x_i\right) = 0$. Puisque A est injective, le lemme 2 dit que $\sum_{i=1}^{k} \lambda_i x_i = 0$, et vu que les vecteurs $x_1, \ldots, x_k$ sont linéairement indépendants, $\lambda_1 = \ldots = \lambda_k = 0$. De ce fait, les vecteurs $Ax_1, \ldots, Ax_k$ sont aussi linéairement indépendants.

2. Opérateurs linéaires inversibles dans $\mathbf{R}^n$

LEMME 4. *Si une application linéaire $A : \mathbf{R}^n \to \mathbf{R}^m$ est injective, $m \geqslant n$.*

Démonstration. Les vecteurs $e^j = (\delta_1^j, \ldots, \delta_n^j)$, $j = 1, \ldots, n$, où δ_k^j est le symbole de Kronecker, forment une base standard de $\mathbf{R}^n$ et sont linéairement indépendants. Si A est injective, les vecteurs Ae^j sont aussi linéairement indépendants en vertu du lemme 3. Ainsi donc, il existe m vecteurs linéairement indépendants dans $\mathbf{R}^m$, d'où il résulte, comme on le sait du cours d'algèbre, que $m \geqslant n$.

COROLLAIRE. *Si une application linéaire $A : \mathbf{R}^n \to \mathbf{R}^m$ est inversible, $m = n$.*

En effet, vu que A et A^{-1} sont injectives dans ce cas, on obtient d'après le lemme 4 que $m \geqslant n$ et $n \geqslant m$.

LEMME 5. *Tout opérateur linéaire injectif A dans $\mathbf{R}^n$ est inversible.*

Démonstration. Comme on l'a vu dans la démonstration du lemme 4, les vecteurs $Ae^1, \ldots, Ae^n$ sont linéairement indépendants. Par conséquent, ils forment une base dans $\mathbf{R}^n$, de sorte que pour tout vecteur $x \in \mathbf{R}^n$ il existe des nombres $\lambda_1, \ldots, \lambda_n$ tels que

$$x = \sum_{j=1}^{n} \lambda_j A e^j = A\left(\sum_{j=1}^{n} \lambda_j e^j\right).$$

Ainsi A est bijective et donc inversible d'après le lemme 1.

THÉORÈME 1. *Si A et B sont des opérateurs linéaires dans $\mathbf{R}^n$ et si de plus A est inversible et $\|B - A\| < \dfrac{1}{\|A^{-1}\|}$, B est également inversible.*

Démonstration. Soit $x \in \mathbf{R}^n$. Comme $x = A^{-1}Bx + A^{-1}(A - B)x$, on a

$$\|x\| \leqslant \|A^{-1}\| \, \|Bx\| + \|A^{-1}\| \, \|A - B\| \, \|x\|,$$

d'où

$$(1 - \|A^{-1}\| \, \|A - B\|)\|x\| \leqslant \|A^{-1}\| \, \|Bx\|. \tag{1}$$

Mais vu que $\|A^{-1}\|\|A - B\| = \|A^{-1}\|\|B - A\| < 1$, on a

$$1 - \|A^{-1}\|\|A - B\| > 0.$$

C'est pourquoi, il résulte de l'inégalité (1) que si $Bx = 0$, i.e. $\|Bx\| = 0$, on a $\|x\| = 0$, i.e. $x = 0$. Il s'ensuit que l'opérateur B est injectif d'après le lemme 2 et inversible suivant le lemme 5.

REMARQUE. Le théorème 1 montre que *les opérateurs linéaires inversibles forment un ensemble ouvert dans l'espace normé* $\mathscr{L}(\mathbf{R}^n)$ *de tous les opérateurs linéaires dans* $\mathbf{R}^n$.

THÉORÈME 2. A^{-1} *est une fonction continue de* A *sur un sous-espace de l'espace normé* $\mathscr{L}(\mathbf{R}^n)$ *formé par les opérateurs linéaires inversibles, i.e.* B *est inversible et* $B^{-1} \to A^{-1}$ *si* A *est inversible et* $B \to A$.

Démonstration. Supposons que A soit inversible. Si $\|B - A\| < \frac{q}{\|A^{-1}\|}$ où $q \in]0, 1[$, l'opérateur B est inversible en vertu du théorème 1. Comme

$$B^{-1} - A^{-1} = A^{-1}(A - B)B^{-1},$$

on a

$$\|B^{-1}\| \leqslant \|A^{-1}\| + \|A^{-1}\|\|A - B\|\|B^{-1}\| < \|A^{-1}\| + q\|B^{-1}\|,$$

d'où $\|B^{-1}\| < \frac{\|A^{-1}\|}{1 - q}$ et alors

$$\|B^{-1} - A^{-1}\| \leqslant \|A^{-1}\|\|A - B\|\|B^{-1}\| \leqslant \frac{\|A^{-1}\|^2}{1 - q}\|A - B\|.$$

Par conséquent, quel que soit $\varepsilon > 0$,

$$\|B - A\| < \min\left(\frac{1 - q}{\|A^{-1}\|}\varepsilon, \frac{q}{\|A^{-1}\|}\right) \Rightarrow \|B^{-1} - A^{-1}\| < \varepsilon,$$

i.e. $B^{-1} \to A^{-1}$ dans $\mathscr{L}(\mathbf{R}^n)$ lorsque $B \to A$.

Rappelons en conclusion qu'un opérateur A dans $\mathbf{R}^n$ est inversible si et seulement si $\det A \neq 0$.

§ 2. Théorèmes des accroissements finis

1. Formule de Lagrange pour les applications de $\mathbf{R}^n$ dans $\mathbf{R}$

Rappelons qu'on appelle *segment* joignant les points a et b dans l'espace $\mathbf{R}^n$ un ensemble

$$[a, b] = \{(1 - t)a + tb \mid 0 \leqslant t \leqslant 1\},$$

i.e. l'image du segment $[0, 1] \subset \mathbf{R}$ par une application $\varphi: \mathbf{R} \to \mathbf{R}^n$ définie suivant la formule

$$\varphi(t) = a + t(b - a). \tag{1}$$

Les points $a = \varphi(0)$ et $b = \varphi(1)$ sont appelés *extrémités* du segment $[a, b]$, et les autres points, *points intérieurs* de ce segment. L'application φ est différentiable (en tant que somme d'une application constante $t \mapsto a$ et d'une application linéaire $t \mapsto t(b - a)$) et

$$\varphi'(t)(\lambda) = \lambda(b - a) \text{ pour tout } t \in \mathbf{R}. \tag{2}$$

Le théorème de Lagrange pour les fonctions de $\mathbf{R}$ dans $\mathbf{R}$ se généralise aisément aux fonctions de $\mathbf{R}^n$ dans $\mathbf{R}$:

THÉORÈME 1. *Si une fonction f de $\mathbf{R}^n$ dans $\mathbf{R}$ est continue aux extrémités du segment $[a, b]$ et différentiable en tous les points intérieurs, on a*

$$f(b) - f(a) = f'(c)(b - a) \tag{3}$$

pour un point intérieur c du segment *) $[a, b]$.

Démonstration. Posons $g = f \circ \varphi$, i.e. $g(t) = f(a + t(b - a))$; g est déjà une fonction de $\mathbf{R}$ dans $\mathbf{R}$. Elle est définie sur le segment $[0, 1]$ et satisfait sur ce segment aux hypothèses du théorème de Lagrange en vertu des théorèmes 2.2.2 et 7.2.4, i.e. elle est continue aux extrémités du segment $[0, 1]$ et dérivable en tous les points intérieurs de celui-ci. En vertu des formules (1) et (2), on a

$$g'(t)(\lambda) = f'(\varphi(t))\varphi'(t)(\lambda) = f'(a + t(b - a))(\lambda(b - a)) = \\ = \lambda f'(a + t(b - a))(b - a) \tag{4}$$

pour tout $t \in [0, 1]$. En appliquant le théorème de Lagrange, on obtient

$$f(b) - f(a) = g(1) - g(0) = g'(\theta)(1) = \\ = f'(a + \theta(b - a))(b - a) = -f'(c)(b - a)$$

où θ est un nombre compris entre 0 et 1, et $c = a + \theta(b - a)$ est un point associé du segment $[a, b]$.

2. Estimation de l'accroissement d'une application différentiable de $\mathbf{R}^n$ dans $\mathbf{R}^m$

Pour les applications de $\mathbf{R}^n$ dans $\mathbf{R}^m$ où $m > 1$, la formule (3) n'est pas en vigueur déjà pour $m = 2$, même si $n = 1$. En effet, supposons que l'application $f: \mathbf{R} \to \mathbf{R}^2$ de composantes f_1 et f_2 soit différentiable sur le

---

*) Ici $f'(c)(b - a) = \sum_{k=1}^{n} f_k'(c)(b_k - a_k)$ où f_k sont les composantes de la fonction f et a_k et b_k, les coordonnées des points a et b.

segment $[a, b] \subset \mathbf{R}$. La formule (3) signifierait dans ce cas qu'il existe un point $c \in]a, b[$ tel que

$$f_i(b) - f_i(a) = \frac{\mathrm{d}f_i}{\mathrm{d}t}(c)(b - a) \qquad (i = 1, 2). \tag{5}$$

Pour chaque composante prise séparément il existe un tel c. Mais il peut arriver que ce point n'est pas le même pour les deux composantes. Si, par exemple,

$$f(t) = (\cos t, \sin t) \text{ et } [a, b] = [0, \pi/2],$$

les formules (5) donneraient

$$\cos\frac{\pi}{2} - \cos 0 = -\sin c\left(\frac{\pi}{2} - 0\right), \text{ i.e. } \sin c = \frac{2}{\pi},$$
$$\sin\frac{\pi}{2} - \sin 0 = \cos c\left(\frac{\pi}{2} - 0\right), \text{ i.e. } \cos c = \frac{2}{\pi},$$

d'où il résulterait que $1 = \frac{8}{\pi^2} < \frac{8}{9}$!

Pour des buts pratiques, il suffit cependant (déjà dans le cas unidimensionnel) que soit vérifiée l'inégalité suivante

$$\|f(b) - f(a)\| \leqslant \sup_{c \in [a,b]} \|f'(c)\| \, \|b - a\|. \tag{6}$$

Par exemple, il en découle immédiatement que f est constante si $f'(x) \equiv 0$. Montrons que l'inégalité (6) a réellement lieu (sous les mêmes hypothèses concernant le comportement de la fonction f sur le segment $[a, b]$ que dans le théorème 1).

LEMME. *Si une application* $g : \mathbf{R}^n \to \mathbf{R}^m$ *est différentiable au point* x_0 *et* v_0 *est un vecteur donné de* $\mathbf{R}^m$, *la fonction* $h(x) = \langle g(x), v_0 \rangle$ *est différentiable au point* x_0 *et*

$$h'(x_0)u = \langle g'(x_0)u, v_0 \rangle \textit{ pour tout } u \in \mathbf{R}^n.$$

Démonstration. Posons $l(v) = \langle v, v_0 \rangle$. Il s'ensuit que l est une fonction linéaire sur $\mathbf{R}^m$ et $h = l \circ g$. Comme g est différentiable au point x_0 et l l'est partout ($l'(v) = l$ pour tout $v \in \mathbf{R}^m$), h est différentiable au point x_0 d'après le théorème 7.2.4, de sorte que

$$h'(x_0) = l'(g(x_0))g'(x_0) = l \circ g'(x_0),$$

i.e. $h'(x_0)u = (l \circ g'(x_0))u = \langle g'(x_0)u, v_0 \rangle$ pour tout $u \in \mathbf{R}^n$.

THÉORÈME 2. *Si une application* $f : \mathbf{R}^n \to \mathbf{R}^m$ *est continue aux extrémités du segment* $[a, b]$ *et différentiable en tous les points intérieurs de ce*

segment,

(7) $$\|f(b) - f(a)\| \leqslant \|f'(c)(b - a)\|$$

pour un point intérieur $c \in [a, b]$.

Démonstration. Tout comme dans la démonstration du théorème 1, supposons que $g = f \circ \varphi$ où φ est une application de $\mathbf{R}$ dans $\mathbf{R}^n$ définie par la formule (1), de sorte que $g(0) = f(a)$, $g(1) = f(b)$. Mais dans ce cas g est une application de $\mathbf{R}$ dans $\mathbf{R}^m$. Pour obtenir une fonction de $\mathbf{R}$ dans $\mathbf{R}$, il suffit de former le produit scalaire de $g(t)$ par un vecteur fixé de $\mathbf{R}^m$. Il est commode de poser

$$h(t) = \langle g(t), f(b) - f(a) \rangle,$$

d'où $h(1) - h(0) = \langle f(b) - f(a), f(b) - f(a) \rangle = \|f(b) - f(a)\|^2$. Tout comme g, la fonction h est continue aux points 0 et 1. Vu que g est différentiable aux points $t \in]0, 1[$ en vertu du théorème 7.2.4, il en est de même pour h d'après le lemme. Ceci étant, on a en vertu de la formule (4) :

$$h'(t)(\lambda) = \langle g'(t)(\lambda), f(b) - f(a) \rangle = \\ = \lambda \langle f'(a + t(b - a))(b - a), f(b) - f(a) \rangle.$$

En appliquant le théorème de Lagrange, on obtient

$$\|f(b) - f(a)\|^2 = h(1) - h(0) = h'(\theta)(1) = \\ = \langle f'(a + \theta(b - a))(b - a), f(b) - f(a) \rangle = \\ = \langle f'(c)(b - a), f(b) - f(a) \rangle$$

où $c = a + \theta(b - a)$ est un point intérieur du segment $[a, b]$. Par conséquent, d'après l'inégalité de Cauchy-Bouniakovski,

$$\|f(b) - f(a)\|^2 \leqslant \|f'(c)(b - a)\| \|f(b) - f(a)\|,$$

d'où, pour $f(a) \neq f(b)$, il résulte l'inégalité (7). Lorsque $f(a) = f(b)$, cette inégalité est triviale.

COROLLAIRE. *Si les hypothèses du théorème 2 sont vérifiées, on a l'inégalité*

$$\|f(b) - f(a)\| \leqslant \|f'(c)\| \|b - a\|$$

où c est un point intérieur de $[a, b]$, et à plus forte raison, l'inégalité (6).

§ 3. Applications différentiables localement inversibles de $\mathbf{R}^n$ dans $\mathbf{R}^n$

1. Critère de différentiabilité de l'application inverse

THÉORÈME 1. *Soit f une application injective de $\mathbf{R}^n$ dans $\mathbf{R}^n$ différentiable au point a (et, par suite, définie dans un voisinage de a). Pour que l'application inverse g soit différentiable au point $b = f(a)$, il est nécessaire et*

suffisant que soient vérifiées les conditions suivantes : 1) *l'opérateur* $f'(a)$ *est inversible*, 2) g *est continue au point* b, 3) $b \in\,]D_g[$. *Si ces conditions sont satisfaites*,

$$g'(b) = [f'(a)]^{-1} \; (= [f'(g(b))]^{-1}). \tag{1}$$

Démonstration. *Nécessité.* La condition (3) est nécessaire d'après la définition de différentiabilité, et la condition 2) l'est puisque la différentiabilité entraîne la continuité. Il reste à démontrer la nécessité de la condition 1). Notons avant tout que I_{D_f} et I_{D_g} sont les restrictions de $I_{\mathbf{R}^n}$ à D_f et D_g respectivement, de sorte que conformément aux exemples 7.2.2 et 7.2.4, I_{D_f} et I_{D_g} sont différentiables respectivement aux points a et b, donc, $I'_{D_f}(a) = I'_{D_g}(b) = I_{\mathbf{R}^n}$. Par conséquent, si g est différentiable au point b, on a $g'(b)f'(a) = I_{\mathbf{R}^n}$, $f'(a)g'(b) = I_{\mathbf{R}^n}$ d'après le théorème 7.2.4, de sorte que l'opérateur $f'(a)$ est inversible et $[f'(a)]^{-1} = g'(b)$.

Suffisance. Il existe pour tout $y \in D_g$ un $x \in D_f$ et un seul tel que $y = f(x)$; ceci étant $x = g(y)$. Par hypothèse, f est différentiable au point a, i.e.

$$y - b = f(x) - f(a) = A(x - a) + \alpha(x)\|x - a\| \tag{2}$$

où $A = f'(a)$ et $\alpha(x) \to 0$ lorsque $x \to a$. D'après la condition (1), A est inversible. En appliquant l'opérateur inverse A^{-1} à deux membres de la formule (2), on obtient, compte tenu de la linéarité de A^{-1},

$$x - a = A^{-1}(y - b) - A^{-1}(\alpha(x))\|x - a\|, \tag{3}$$

i.e.

$$g(y) - g(b) = A^{-1}(y - b) + \beta(y)\|y - b\| \tag{4}$$

où

$$\beta(y) = \begin{cases} -A^{-1}(\alpha(g(y)))\dfrac{\|g(y) - g(b)\|}{\|y - b\|} & \text{lorsque } y \neq b, \\ 0 & \text{lorsque } y = b. \end{cases}$$

Montrons que $\beta(y) \to 0$ lorsque $y \to b$. D'après la condition 2), $g(y) \to$ $\to g(b)$ lorsque $y \to b$; comme $\alpha(x) \to 0$ lorsque $x \to a$, on a $\alpha(g(y)) \to 0$ lorsque $y \to b$; il résulte de la linéarité de l'opérateur A^{-1} qu'il est continu et $A^{-1}(0) = 0$, si bien que $-A^{-1}(\alpha(g(y))) \to -A^{-1}(0) = 0$ lorsque $y \to$ $\to b$ d'après le théorème de passage à la limite sous le signe d'une fonction continue. Il reste à montrer qu'il existe un voisinage de b dans lequel $\dfrac{\|g(y) - g(b)\|}{\|y - b\|}$ est bornée. Il découle de la formule (3) que

$$\|x - a\| \leqslant \|A^{-1}\|\,\|y - b\| + \|A^{-1}(\alpha(x))\|\,\|x - a\|,$$

i.e.

$$(1 - \|A^{-1}(\alpha(g(y)))\|)\|g(y) - g(b)\| \leqslant \|A^{-1}\| \|y - b\|. \tag{5}$$

Or $A^{-1}(\alpha(g(y))) \to 0$ quand $y \to b$, et $b \in]D_g[$, de sorte que D_g contient un voisinage V du point b tel que

$$\|A^{-1}(\alpha(g(y)))\| < \frac{1}{2} \text{ pour tout } y \in V.$$

Il résulte alors de l'inégalité (5) que

$$\frac{\|g(y) - g(b)\|}{\|y - b\|} \leqslant \frac{\|A^{-1}\|}{1 - \|A^{-1}(\alpha(g(y)))\|} < 2\|A^{-1}\|$$

pour tout $y \in V \setminus \{b\}$. Enfin, puisque $b \in]D_g[$, A^{-1} est un opérateur linéaire et $\beta(y) \to 0$ lorsque $y \to b$; la formule (4) montre que g est différentiable au point b et la formule (1) est vraie.

2. Conditions suffisantes pour qu'une application soit localement injective et qu'elle soit un homéomorphisme

Les conditions nécessaires et suffisantes du théorème 1 sont imposées non seulement à f mais aussi à g. Il serait souhaitable d'avoir au moins des conditions suffisantes qui soient imposées à f seule. Ce théorème sera obtenu au n° 4. Pour abréger sa démonstration, nous en prendrons deux fragments qui présentent un intérêt à soi et qui assurent la vérification des conditions du théorème 1 ; le présent numéro et les numéros qui suivent y seront consacrés.

THÉORÈME 2. *Soit f une application différentiable de* $\mathbf{R}^n$ *dans* $\mathbf{R}^n$. *Si l'opérateur $A = f'(a)$ est inversible et la dérivée f' de l'application f est continue au point a, i.e. $\|f'(x) - f'(a)\| \to 0$ lorsque $x \to a$, il existe un voisinage ouvert U de a contenu dans D_f tel que l'opérateur $f'(x)$ est inversible pour tout $x \in U$, la restriction $f|_U$ de f à U est injective et l'application g inverse à $f|_U$ est continue.*

Démonstration. Choisissons un $q \in]0, 1[$ arbitraire. Comme A est inversible et f' est continue au point a, il existe une boule ouverte U de centre a contenue dans D_f telle que

$$\|f'(x) - A\| < \frac{q}{\|A^{-1}\|} \text{ pour tout } x \in U. \tag{6}$$

Puisque $q < 1$, *l'opérateur $f'(x)$ est inversible pour tout $x \in U$* d'après le théorème 1 du § 1. Montrons que U possède les deux autres propriétés affirmées par le théorème.

Posons $\varphi = f - A$. Comme A est linéaire, φ est différentiable et $\varphi'(x) = f'(x) - A'(x) = f'(x) - A$. Ainsi donc,

$$(6') \qquad \|\varphi'(x)\| < \frac{q}{\|A^{-1}\|} \text{ pour tout } x \in U.$$

Soient $x_1, x_2 \in U$ et $x_1 \neq x_2$. On a vu au n° 3.1.1 que U est convexe, de sorte que $[x_1, x_2] \subset U$. En vertu du corollaire au théorème 2 du § 2, on a

$$\|\varphi(x_2) - \varphi(x_1)\| \leqslant \|\varphi'(c)\| \|x_2 - x_1\|$$

où c est un point intérieur du segment $[x_1, x_2]$, appartenant par là même à U. C'est pourquoi, en vertu de (6'),

$$(7) \quad \|f(x_2) - f(x_1) - A(x_2 - x_1)\| = \|\varphi(x_2) - \varphi(x_1)\| < \frac{q}{\|A^{-1}\|} \|x_2 - x_1\|.$$

D'autre part, vu que

$$\|x_2 - x_1\| = \|A^{-1}A(x_2 - x_1)\| = \|A^{-1}\| \|A(x_2 - x_1)\|,$$

on a

$$(8) \qquad \|A(x_2 - x_1)\| \geqslant \frac{1}{\|A^{-1}\|} \|x_2 - x_1\|.$$

Puisque

$$\|f(x_2) - f(x_1)\| \geqslant \|A(x_2 - x_1)\| - \|f(x_2) - f(x_1) - A(x_2 - x_1)\|,$$

il résulte des inégalités (7) et (8) que

$$(9) \qquad \|f(x_2) - f(x_1)\| \geqslant \frac{1 - q}{\|A^{-1}\|} \|x_2 - x_1\|$$

pour tous $x_1, x_2 \in U$ (l'égalité n'est obtenue que si $x_1 = x_2$). L'inégalité (9) montre avant tout (puisque $1 - q > 0$) que $f(x_1) \neq f(x_2)$ si $x_1, x_2 \in U$ et $x_1 \neq x_2$, i.e. *la restriction* $f|_U$ *de l'application* f *à* U *est injective* et, par suite, possède une application inverse g. Soient $y_1, y_2 \in D_g$ $(= f(U))$. Posons $x_1 = g(y_1)$, $x_2 = g(y_2)$, de sorte que $y_1 = f(x_1)$, $y_2 = f(x_2)$. L'inégalité (9) se met alors sous la forme suivante :

$$(10) \qquad \|g(y_2) - g(y_1)\| \leqslant \frac{\|A^{-1}\|}{1 - q} \|y_2 - y_1\|$$

pour tous $y_1, y_2 \in D_g$. Il résulte de l'inégalité (10) que $g(y_2) \to g(y_1)$ lorsque $y_2 \to y_1$, i.e. *g est continue en chaque point $y_1 \in D_g$.*

3. Image ouverte

L'application continue f possédant un inverse continu est appelée *homéomorphisme de D_f sur $f(D_f)$*. Ainsi donc, *si les hypothèses du théorème* 2 *sont vérifiées,* $f|_U$ (étant continue avec f et possédant un inverse continu) *est un homéomorphisme de U sur $f(U)$*. Comme U est un ensemble ouvert

dans $\mathbf{R}^n$ et $f(U) \subset \mathbf{R}^n$, on obtient d'après le théorème de Brouwer (topologie) que $f(U)$ est un ensemble ouvert dans $\mathbf{R}^n$. Mais la démonstration du théorème de Brouwer dépasse le cadre du cours d'Analyse mathématique. Pour démontrer que $f(U)$ est ouvert, on utilisera dans ce numéro les propriétés différentielles de l'application f.

THÉORÈME 3. *Soient f une application différentiable de $\mathbf{R}^n$ dans $\mathbf{R}^n$ et U un ensemble ouvert contenu dans D_f. Si la restriction $f|_U$ de f à U est injective et l'opérateur $f'(x)$ est inversible pour tout $x \in U$, $V = f(U)$ est un ouvert.*

Démonstration. Soit $y^0 \in V$, de sorte que $y^0 = f(x^0)$ où $x^0 \in U$. Etant un voisinage du point x^0, U contient une boule fermée *) $B(x^0\,;r)$. En effet, U contient un ε-voisinage $U(x^0\,;\varepsilon)$ du point x^0 et $B(x^0\,;r) \subset \subset U(x^0\,;\varepsilon)$ pour tout $r \in]0, \varepsilon[$. Vu que $x^0 \notin S(x^0\,;r)$ et $f|_U$ est injective, $y^0 \notin f(S(x^0\,;r))$. Soit ϱ la distance de y^0 à $f(S(x^0\,;r))$ (voir n° 3.2.5). Puisque $S(x^0\,;r)$ est un compact (voir exemple 3.2.3) et f est continue parce que différentiable, $f(S(x^0\,;r))$ est un compact et donc $\varrho > 0$ selon la propriété 3.2.5, 2°. Montrons que $U\left(y^0\,;\frac{\varrho}{2}\right) \subset f(U(x^0\,;r))$. Soit $y^1 \in U\left(y^0\,;\frac{\varrho}{2}\right)$, i.e. $\|y^1 - y^0\| < \frac{\varrho}{2}$. Comme $B(x^0\,;r)$ est un compact (voir exemple 3.2.1) et $\|y^1 - f(x)\|$ est une fonction continue, celle-ci présente en un point $x^1 \in B(x^0\,;r)$ sa valeur minimale. On a $x^1 \notin S(x^0\,;r)$ car, si $x \in S(x^0\,;r)$, on aurait

$$\|y^1 - f(x)\| \geqslant \|y^0 - f(x)\| - \|y^1 - y^0\| > \varrho - \frac{\varrho}{2} = \frac{\varrho}{2},$$

tandis que $\|y^1 - f(x^1)\| \leqslant \|y^1 - f(x^0)\| = \|y^1 - y^0\| < \frac{\varrho}{2}$. Ainsi donc, $x^1 \in U(x^0\,;r)$, de sorte que $\|y^1 - f(x^0)\|$ prend sa valeur minimale sur la boule $B(x^0\,;r)$ en un point intérieur x^1 de celle-ci. La même boule contient aussi le point de minimum de

$$\|y^1 - f(x)\|^2 = \sum_{i=1}^{n}[y_i^1 - f_i(x)]^2 \tag{11}$$

où y_i^1 sont les coordonnées du point y^1 et f_i, les composantes de l'application f. Vu que (11) est une fonction différentiable, ses dérivées partielles d'ordre 1 s'annulent au point x^1 d'après le théorème 8.3.1. Or

$$\frac{\partial}{\partial x_j}\|y^1 - f(x)\|^2 = -2\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_j}(x)[y_i^1 - f_i(x)].$$

---

*) Rappelons que $B(x^0\,;r) = \{x \mid d(x^0, x) \leqslant r\}$, $U(x^0\,;r) = \{x \mid d(x^0, x) < r\}$; $S(x^0\,;r) = \{x \mid d(x^0, x) = r\}$, de sorte que $B(x^0\,;r) = U(x^0\,;r) \cup S(x^0\,;r)$.

Ainsi donc,

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(x^1)[y_i^1 - f_i(x^1)] = 0 \qquad (j = 1, \ldots, n).$$

C'est un système de n équations linéaires homogènes en $y_i^1 - f_i(x^1)$ $(i = 1, \ldots, n)$ dont le déterminant est le jacobien $\det f'(x^1)$ de l'application f au point x^1 (voir n° 7.2.3). Comme $x^1 \in U$, l'opérateur $f'(x^1)$ est inversible, i.e. $\det f'(x^1) \neq 0$. Par conséquent,

$$y_i^1 - f_i(x^1) = 0 \qquad (i = 1, \ldots, n),$$

i.e. $y^1 = f(x^1) \in f(U(x^0 ; r))$. Cela démontre que $U\left(y^0 ; \frac{\varrho}{2}\right) \subset$ $\subset f(U(x^0 ; r)) \subset f(U) = V$. Donc, tout point y de V est contenu dans V avec l'un quelconque de ses voisinages, i.e. V est un ouvert.

4. Théorème définitif

THÉORÈME 4. *Soit f une application différentiable de $\mathbf{R}^n$ dans $\mathbf{R}^n$. Si sa dérivée f' est continue au point a et l'opérateur $A = f'(a)$ est inversible, il existe des voisinages ouverts $U \subset D_f$ et V des points respectifs a et $b = f(a)$ tels que l'opérateur $f'(x)$ est inversible pour tout $x \in U$, la restriction de f à U est un homéomorphisme de U sur V et l'application g inverse à $f|_U$ est différentiable. Ceci étant,*

(12) $$g'(y) = [f'(g(y))]^{-1} \text{ pour tout } y \in V$$

et g' est continue au point $y = b$.

Démonstration. Les hypothèses du théorème 2 étant vérifiées, il existe un voisinage ouvert U du point a contenu dans D_f tel que l'opérateur $f'(x)$ est inversible pour tout $x \in U$ et la restriction $f|_U$ est injective. Ainsi donc, les hypothèses du théorème 3 sont vérifiées et $V = f(U)$ est un ouvert. En outre, d'après le théorème 2, l'application g inverse à $f|_U$ est continue en tout point $y \in V$, donc $f|_U$ est un homéomorphisme de U sur V. Vu que U est un ouvert, on obtient conformément à l'exemple 7.2.4 que $f|_U$ et f sont différentiables en tout point $x \in U$ et $(f|_U)'(x) = f'(x)$. Ainsi donc, l'application $f|_U$ (au lieu de f), tout point $x \in U$ (au lieu de a) et $y = f(x)$ (au lieu de b) vérifient toutes les conditions du théorème 1 : l'application $f|_U$ est injective et est différentiable au point x, l'opérateur $(A_U)'(x)$ est inversible, g est continue au point $y = (f|_U)(x)$ et $y \in]D_g[\;(= V)$. Par conséquent, g est différentiable en y et $g'(y) = [(f|_U)'(g(y))]^{-1} =$ $= [f'(g(y))]^{-1}$. Enfin, g est continue au point $y = b$, $f'(x)$ l'est au point $x = a = g(b)$ et, d'après le théorème 2 du § 1, A^{-1} dépend continûment de A. Il résulte alors de la dernière formule que g' est continue au point $y = b$.

REMARQUE. Si une application $f: \mathbf{R}^n \to \mathbf{R}^n$ est continûment différentiable (i.e. f est différentiable et f', continue) et l'opérateur $f'(x)$ est inversible pour tout $x \in D_f$, il découle du théorème 2 que f est « localement injective », i.e. tout point $x \in D_f$ possède un voisinage ouvert $U \in D_f$ tel que $f|_U$ est injective. Cependant, f n'est pas obligatoirement injective. Ainsi, une application $f: \mathbf{R}^2 \to \mathbf{R}^2$ définie par

$$D_f = \{(r, \varphi) \in \mathbf{R}^2 \mid r > 0\} \text{ et } f(r, \varphi) = (r \cos \varphi, r \sin \varphi)$$

(voir exemple 7.2.5) est continûment différentiable (voir plus loin théorème 5) et l'opérateur $f'(r, \varphi)$ est inversible pour tout $(r, \varphi) \in D_f$ puisque $\det f'(r, \varphi)(= r)$ n'est pas nul. Cependant, $f(r, \varphi) = f(r, \varphi + 2\pi)$, de sorte que f n'est pas injective.

5. Critère de différentiabilité continue

Dans les théorèmes 2 et 4 un rôle important appartient à la condition de la continuité de la dérivée de l'application f au point donné a, plus précisément à la condition $\|f'(x) - f'(a)\| \to 0$ lorsque $x \to a$. On peut formuler cette condition en d'autres termes sans recourir à la notion de norme de l'application linéaire.

LEMME. *Soit A une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et soit (A_{ij}) sa matrice dans les bases standard. Alors*

$$|A_{ij}| \leqslant \|A\| \leqslant \sqrt{\sum_{j=1}^{n} \sum_{i=1}^{m} A_{ij}^2}.$$

Démonstration. Comme Ae^j est la j-ième colonne de la matrice (A_{ij}) (voir n° 7.1.2), $\|Ae^j\|^2 = \sum_{i=1}^{m} A_{ij}^2$ et, par suite, $|A_{ij}| \leqslant \|Ae^j\| \leqslant \|A\| \|e^j\| = \|A\|$. Vu que

$$\|Ax\| \leqslant \sqrt{\sum_{j=1}^{n} \|Ae^j\|^2} \|x\| \text{ pour tout } x \in \mathbf{R}^n$$

(formule 7.1(1)), on a $\|A\| \leqslant \sqrt{\sum_{j=1}^{n} \|Ae^j\|^2} = \sqrt{\sum_{j=1}^{n} \sum_{i=1}^{m} A_{ij}^2}$.

THÉORÈME 5. *Soit f une application de $\mathbf{R}^n$ dans $\mathbf{R}^m$ différentiable au point a et soient $f_1, \ldots, f_m$ ses composantes ; f' est continue au point a si et seulement s'il en est de même pour toutes les dérivées partielles $\frac{\partial f_i}{\partial x_j}$* $(i = 1, \ldots, m ; j = 1, \ldots, n)$.

Démonstration. En vertu du théorème 7.2.6, la matrice de l'application linéaire $f'(x) - f'(a)$ est

$$\left(\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right).$$

Il en découle d'après le lemme que pour tout entier $i \in [1, m]$ et tout entier $j \in [1, n]$ on a les inégalités suivantes

$$\left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right| \leqslant \|f'(x) - f'(a)\| \leqslant \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right|^2}.$$

La première inégalité montre que si $\|f'(x) - f'(a)\| \to 0$ lorsque $x \to a$, on a $\frac{\partial f_i}{\partial x_j}(x) \to \frac{\partial f_i}{\partial x_j}(a)$ lorsque $x \to a$ pour tous i et j, et la deuxième montre que la réciproque est aussi vraie.

CHAPITRE 10

FONCTIONS IMPLICITES ET EXTRÉMUMS RELATIFS

§ 1. Fonctions implicites

Le terme plus précis serait « fonctions implicitement définies ». Il s'agit des fonctions de $\mathbf{R}^n$ dans $\mathbf{R}^m$ satisfaisant à une équation non résolue. Sous quelles conditions (et où) ces fonctions existent-elles ? Sont-elles différentiables, si oui, comment calculer leur dérivée ?

1. Localisation du problème

Prenons, pour commencer, le cas unidimensionnel. Soit une équation

$$f(x, y) = 0 \tag{1}$$

où x et y sont des variables réelles et f est une fonction réelle. Est-ce qu'il existe (et où) une fonction $g(x)$ dont la substitution à y transforme cette équation en une identité (i.e. $f(x, g(x)) = 0$ pour tout $x \in D_g$) ? Quelles sont ses propriétés ? Une telle fonction peut ne pas exister du tout. Ainsi, il n'existe pas de nombres réels x et y qui vérifient l'équation

$$x^2 + y^2 + 1 = 0.$$

La fonction $g(x)$ peut ne pas être unique. Ainsi, l'équation

$$x^2 + y^2 - 1 = 0 \tag{2}$$

possède deux solutions évidentes : $y = \sqrt{1 - x^2}$ et $y = -\sqrt{1 - x^2}$ (et un ensemble infini d'autres solutions !). Cependant, dans cet exemple, on peut obtenir une solution unique en localisant son étude dans le voisinage requis d'un point (a, b) qui satisfait à l'équation (2) et qui n'appartient pas à l'axe des abscisses : un voisinage suffisamment petit d'un tel point intercepte, sur la circonférence (2), le graphe d'une fonction de x (fig. 27). Mais cela ne réussit pas pour les points $(\pm 1, 0)$, quelque petit que soit leur voisinage (fig. 28). Il en est de même pour l'équation $x^2 - y^2 = 0$ où le point exclusif est l'origine des coordonnées (fig. 29).

Ainsi donc, il surgit le problème « local » suivant. Etant donné un point (a, b) satisfaisant à l'équation (1), trouver les conditions imposées à f sous lesquelles il existe un voisinage de (a, b) dont tous les points (x, y) qui vérifient l'équation (1), forment le graphe d'une fonction de x.

Le problème analogue surgit dans le cas multidimensionnel, à savoir quand $x \in \mathbf{R}^n$ et $y \in \mathbf{R}^m$ dans l'équation (1), de sorte que cette équation

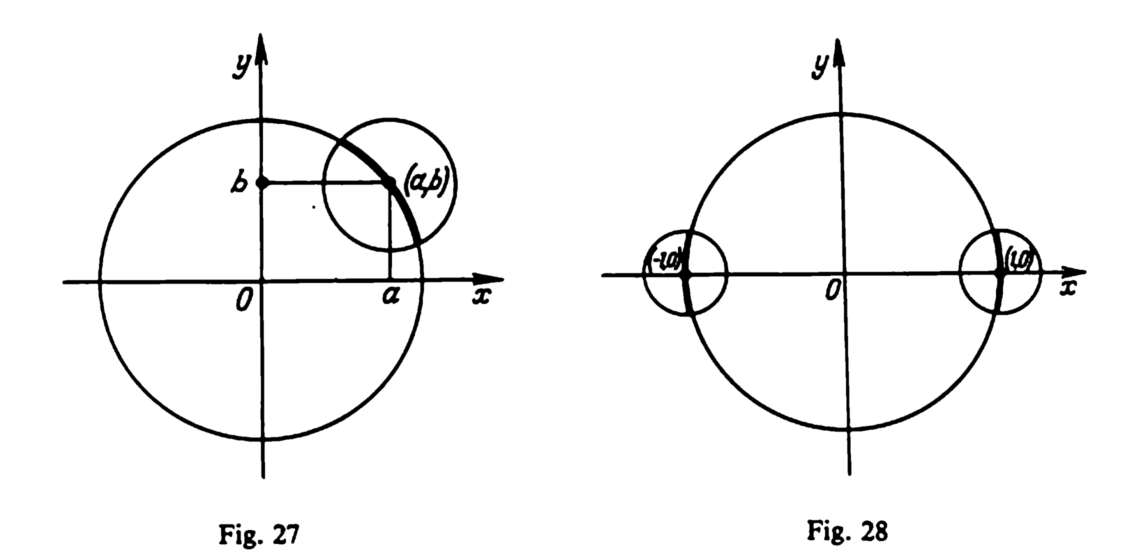


Fig. 27 Fig. 28

s'écrit en coordonnées sous la forme suivante :

$$f(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0.$$

Quand cette équation est-elle résoluble par rapport à $y_1, \ldots, y_m$? On sait du cours d'Algèbre élémentaire que pour définir m nombres inconnus, il faut m équations, i.e. il faut dans ce cas que f soit une fonction de $\mathbf{R}^{n+m}$ dans $\mathbf{R}^m$. Soient $f_1, \ldots, f_m$ ses composantes. L'équation (1) est alors une notation abrégée d'un système de m équations

$$(1') \qquad f_i(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0 \quad (i = 1, \ldots, m)$$

par rapport à m inconnus $y_1, \ldots, y_m$. Le problème local est posé comme suit. Etant donné un point $(a, b) = (a_1, \ldots, a_n, b_1, \ldots, b_m)$ satisfaisant à l'équation (1), i.e. au système (1 '), déterminer les conditions sous lesquel-

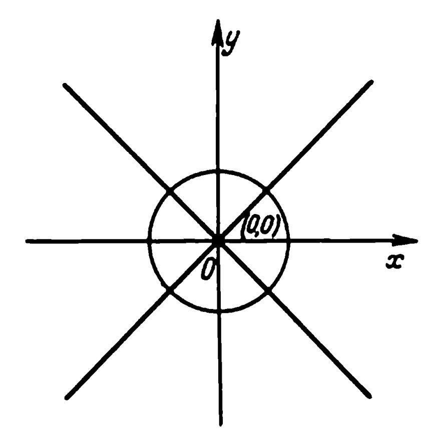


Fig. 29

les il existe un voisinage V de ce point tel que $V \cap K$, où

$$K = \{(x, y) \in \mathbf{R}^{n+m} \mid f(x, y) = 0\},$$

soit le graphe d'une application g de $\mathbf{R}^n$ dans $\mathbf{R}^m$, i.e.

$$(x, y) \in V \cap K \Leftrightarrow x \in D_g \text{ et } y = g(x).$$

2. Conditions suffisantes de la solubilité locale

Les points de l'espace $\mathbf{R}^{n+m}$ seront présentés sous la forme

$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m)$$

où $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ et $y = (y_1, \ldots, y_m) \in \mathbf{R}^m$.

THÉORÈME 1. *Soit f une application différentiable de $\mathbf{R}^{n+m}$ dans $\mathbf{R}^m$ satisfaisant aux conditions suivantes :*

1) $f(a, b) = 0$,

2) *f' est continue au point (a, b),*

3) *le déterminant*

$$J(x, y) = \begin{vmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \ldots & \frac{\partial f_1}{\partial y_m}(x, y) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial y_1}(x, y) & \ldots & \frac{\partial f_m}{\partial y_m}(x, y) \end{vmatrix}$$

est non nul au point (a, b).

Il existe alors des voisinages ouverts U et V des points respectifs $a \in \mathbf{R}^n$ et $(a, b) \in \mathbf{R}^{n+m}$, et une application $g : \mathbf{R}^n \to \mathbf{R}^m$, avec $D_g = U$, tels que $V \subset D_f$ et

$$\{(x, y) \in V \mid f(x, y) = 0\} = \{(x, y) \in \mathbf{R}^{n+m} \mid x \in U \text{ et } y = g(x)\},$$

i.e. $V \cap K$ est le graphe de g. Ceci étant, g est différentiable et g' est continue au point a.

Démonstration. Introduisons une application auxiliaire F de $\mathbf{R}^{n+m}$ dans $\mathbf{R}^{n+m}$ définie par les formules

$$D_F = D_f \text{ et } F(x, y) = (x, f(x, y)) \text{ pour tout } (x, y) \in D_F.$$

En vertu de la condition 1) du théorème,

$$F(a, b) = (a, 0).$$

L'application F est la somme des applications $(x, y) \mapsto (x, 0)$ et $(x, y) \mapsto (0, f(x, y))$ dont la première est linéaire et, pour cette raison, partout différentiable et la deuxième est la composée de l'application différentiable f et de l'application linéaire $v \mapsto (0, v)$ de $\mathbf{R}^m$ dans $\mathbf{R}^{n+m}$. Par conséquent,

F est différentiable. Cherchons sa matrice jacobienne. Soient $F_1, \ldots, F_{n+m}$ les composantes de l'application F, de sorte que

$$F_k(x, y) = \begin{cases} x_k & (1 \leqslant k \leqslant n), \\ f_{k-n}(x, y) & (n+1 \leqslant k \leqslant n+m). \end{cases}$$

Alors

$$(3) \qquad \frac{\partial F_k}{\partial x_j}(x, y) = \begin{cases} \delta_k^j & (1 \leqslant k \leqslant n), \\ \dfrac{\partial f_{k-n}}{\partial x_j}(x, y) & (n+1 \leqslant k \leqslant n+m) \end{cases}$$

(où δ_k^j est le symbole de Kronecker) et

$$(4) \qquad \frac{\partial F_k}{\partial y_i}(x, y) = \begin{cases} 0 & (1 \leqslant k \leqslant n), \\ \dfrac{\partial f_{k-n}}{\partial y_i}(x, y) & (n+1 \leqslant k \leqslant n+m). \end{cases}$$

Par suite, la matrice jacobienne de l'application F au point (x, y) est de la forme

$$\left[\begin{array}{ccc|ccc} 1 & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & 1 & & & \\ \hline \dfrac{\partial f_1}{\partial x_1}(x, y) & \ldots & \dfrac{\partial f_1}{\partial x_n}(x, y) & \dfrac{\partial f_1}{\partial y_1}(x, y) & \ldots & \dfrac{\partial f_1}{\partial y_m}(x, y) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \dfrac{\partial f_m}{\partial x_1}(x, y) & \ldots & \dfrac{\partial f_m}{\partial x_n}(x, y) & \dfrac{\partial f_m}{\partial y_1}(x, y) & \ldots & \dfrac{\partial f_m}{\partial y_m}(x, y) \end{array}\right]$$

Le déterminant de cette matrice, i.e. le jacobien det $F'(x, y)$ de l'application F au point (x, y), est égal à $J(x, y)$. Ainsi donc, det $F'(a, b) \neq 0$ en vertu de la condition 3) du théorème, i.e. l'opérateur $F'(a, b)$ est inversible. En outre, f' est continue au point (a, b) d'après la condition 2), si bien que toutes les dérivées partielles $\dfrac{\partial f_l}{\partial x_j}$ et $\dfrac{\partial f_l}{\partial y_i}$ sont continues en ce point d'après le théorème 9.3.5. Il s'ensuit en vertu des formules (3) et (4) que toutes les dérivées partielles $\dfrac{\partial F_k}{\partial x_j}$ et $\dfrac{\partial F_k}{\partial y_i}$ y sont également continues, de sorte que, d'après le même théorème, F' est continue au point (a, b).

Ainsi donc, F vérifie les hypothèses du théorème 9.3.4 : elle est différentiable, sa dérivée F' est continue au point (a, b) et l'opérateur $F'(a, b)$ est inversible. Par conséquent, en vertu de ce théorème, il existe des voisinages ouverts $V \subset D_F$ et W des points respectifs (a, b) et $F(a, b) = (a, 0)$ tels que l'opérateur $F'(x, y)$ est inversible pour tout $(x, y) \in V$, la restriction de F à V est un homéomorphisme de V sur W et l'application G inverse à la restriction de F à V est différentiable, G' étant continue au point $(a, 0)$. Soit l une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^{n+m}$ définie par la formule

$$l(x) = (x, 0) \text{ pour tout } x \in \mathbf{R}^n$$

et soit $U = l^{-1}(W)$. Vu que l est continue, U et W sont des ouverts en vertu du théorème 2.2.3. Ceci étant, $a \in U$ puisque $l(a) = (a, 0) \in W$. Donc, U est un voisinage ouvert du point a. Soient π et σ des projecteurs de $\mathbf{R}^{n+m}$ sur $\mathbf{R}^m$ et $\mathbf{R}^n$ respectivement, i.e. $\pi(x, y) = y$ et $\sigma(x, y) = x$ pour tout $(x, y) \in \mathbf{R}^{n+m}$. Posons

$$g = \pi \circ G \circ l \qquad \text{et} \qquad h = \sigma \circ G \circ l.$$

Comme π est partout défini, $D_g = D_{G \circ l} = \{x \in \mathbf{R}^n \mid l(x) = D_G\} = U$ puisque $D_g = W$. De façon analogue, $D_h = U$. Ainsi donc, si $x \in U$, i.e. $(x, 0) \in W$, on a

$$G(x, 0) = (G \circ l)(x) = \sigma((G \circ l)(x)), \qquad \pi((G \circ l)(x)) = (h(x), g(x)).$$

On obtient alors

$$(x, 0) = F(G(x, 0)) = F(h(x), g(x)) = (h(x), f(h(x), g(x))),$$

d'où premièrement $h(x) = x$, de sorte que

$$G(x, 0) = (x, g(x)) \text{ pour tout } x \in U \tag{5}$$

(et donc $(x, g(x)) \in V$), et deuxièmement, $f(x, g(x)) = 0$, i.e. $(x, g(x)) \in K$. Ainsi donc, $(x, g(x)) \in V \cap K$ pour tout $x \in U$. Inversement : si $(x, y) \in$ $\in V \cap K$, on obtient en vertu de la formule (5) que

$$(x, y) = G(F(x, y)) = G(x, f(x, y)) = G(x, 0) = (x, g(x)),$$

de sorte que $x \in U$ et $y = g(x)$. Notons que $b = g(a)$ puisqu'en particulier $(a, b) \in V \cap K$.

La différentiabilité de g résulte de la formule $g = \pi \circ G \circ l$ en vertu du théorème 7.2.4 puisque l et π sont différentiables en tant qu'applications linéaires et G l'est d'après ce qu'il a été dit plus haut. Donc,

$$g'(x) = (\pi \circ G)'(l(x))l'(x) = \pi'(G(l(x))) \circ G'(l(x)) \circ l'(x),$$

i.e.

$$g'(x) = \pi \circ G'(l(x)) \circ l = \pi \circ G'(x, 0) \circ l \text{ pour tout } x \in U. \tag{6}$$

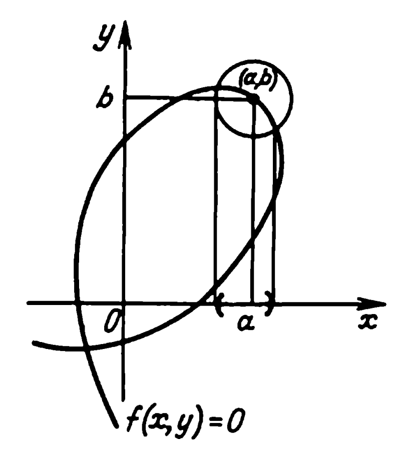


Fig. 30

Enfin, la continuité de g' au point a résulte de la formule (6) en vertu du théorème 2.2.2 puisque l et π sont partout continus, et G' est continue au point $(a, 0) = l(a)$ d'après ce qu'il a été dit plus haut.

Le théorème 1 est ainsi démontré. Pour $m = n = 1$ il prend la forme suivante (compte tenu du théorème 9.3.5).

THÉORÈME 1′. *Soit f une fonction réelle différentiable de deux variables réelles, qui vérifie les conditions suivantes* :

1) $f(a, b) = 0$,

2) $\dfrac{\partial f}{\partial x}$ *et* $\dfrac{\partial f}{\partial y}$ *sont continues au point* (a, b),

3) $\dfrac{\partial f}{\partial y}(a, b) \neq 0$.

Alors D_f contient un voisinage ouvert du point (a, b) interceptant, sur la courbe $f(x, y) = 0$, le graphe d'une fonction dérivable $g(x)$ définie dans un voisinage ouvert du point a (fig. 30). *Ceci étant, g' est continue au point a.*

Ainsi la fonction $f(x, y) = x^2 + y^2 - 1$ est différentiable et possède des dérivées partielles continues. Par ailleurs, $f'_y(x, y) = 2y$. C'est pourquoi, les conditions du théorème 1′ sont remplies pour tout point (a, b) satisfaisant à l'équation (2) et tel que $b \neq 0$, i.e. $a \neq \pm 1$. Comme nous l'avons vu, la conclusion du théorème n'est plus en vigueur aux points $(\pm 1, 0)$.

3. Calcul de la dérivée

Soient f une application de $\mathbf{R}^{n+m}$ dans $\mathbf{R}^m$, (a, b), un point de D_f vérifiant les conditions du théorème 1 et g une application de $\mathbf{R}^n$ dans $\mathbf{R}^m$ dont l'existence et les propriétés sont affirmées par ce théorème. En particulier, $f(x, g(x)) = 0$ sur U, ou en notation analytique :

$$f_i(x_1, \ldots, x_n, g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)) \equiv 0$$
$$(i = 1, \ldots, m)$$

où f_i sont les composantes de l'application f et $g_1, \ldots, g_m$, celles de l'application g. Vu que les fonctions f_i et g_k sont différentiables avec f et g, le théorème 9.2.9 implique pour $y = g(x)$:

$$(7) \quad \frac{\partial f_i}{\partial x_j}(x, y) + \frac{\partial f_i}{\partial y_1}(x, y)\frac{\partial g_1}{\partial x_j}(x) + \ldots + \frac{\partial f_i}{\partial y_m}(x, y)\frac{\partial g_m}{\partial x_j}(x) = 0$$
$$(i = 1, \ldots, m\,;\, j = 1, \ldots, n).$$

Pour tout entier $j \in [1, n]$ c'est un système de m équations du premier degré à m inconnues $\frac{\partial g_1}{\partial x_j}(x), \ldots, \frac{\partial g_m}{\partial x_j}(x)$ (où $x \in U$). Son déterminant (pour $y = g(x)$) est $J(x, g(x)) = \det F'(x, g(x))$. Il découle de $(x, g(x)) \in V$ que $F'(x, g(x))$ est inversible et donc $J(x, g(x)) \neq 0$. Par conséquent, le système (7) est univoquement soluble pour tout $j \in [1, n]$. Ayant résolu ces n systèmes, on obtient les dérivées de toutes les composantes de l'application g :

$$g_i'(x)(h) = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(x)h_j \qquad (i = 1, \ldots, m)$$

et la matrice jacobienne de celle-ci

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_n}(x) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \frac{\partial g_m}{\partial x_1}(x) & \cdots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix}.$$

Dans le cas particulier où $m = 1$, les équations (7) prennent la forme

$$\frac{\partial f}{\partial x_j}(x, y) + \frac{\partial f}{\partial y}(x, y)\frac{\partial g}{\partial x_j}(x) = 0,$$

d'où

$$\frac{\partial g}{\partial x_j}(x) = -\frac{\partial f}{\partial x_j}(x, y)\Big/\frac{\partial f}{\partial y}(x, y) \qquad (j = 1, \ldots, n).$$

Il est évident que les dérivées partielles $\frac{\partial g}{\partial x_j}(x)$ sont données implicitement si l'équation $f(x, y) = 0$ n'est pas résolue.

§ 2. Extrémums liés

1. Notions de maximums et de minimums liés

L'*extrémum lié* d'une fonction réelle de plusieurs variables réelles est son maximum ou minimum si l'on tient compte des seuls points de son ensemble de définition dont les coordonnées sont liées par une ou par plusieurs équations données (dites « contraintes égalité »). Ainsi, la fonction $\varphi(x, y) = xy$ n'a pas d'extrémums. Cependant, elle possède un minimum au point (0, 0) sous la condition supplémentaire $y = x$ (i.e. sur la bissectrice du premier et du troisième quadrant) et un maximum au même point sous la condition $y = -x$.

En général, étant donné une fonction $\varphi : \mathbf{R}^{n+m} \to \mathbf{R}$ et une application $f : \mathbf{R}^{n+m} \to \mathbf{R}^m$ de composantes $f_1, \ldots, f_m$, on dit que le point $(a, b) = (a_1, \ldots, a_n, b_1, \ldots, b_m) \in D_\varphi \cap D_f$ est un *point de maximum lié* (resp. *de minimum lié*) de la fonction φ sous les contraintes $f_1(x, y) = 0, \ldots \ldots, f_m(x, y) = 0$ s'il satisfait à ces équations, i.e. $f(a, b) = 0$, et s'il possède un voisinage ouvert $Q \subset D_\varphi \cap D_f$ tel que $\varphi(a, b) \geqslant \varphi(x, y)$ (resp. $\varphi(a, b) \leqslant \varphi(x, y)$) pour tout $(x, y) \in Q$ vérifiant l'équation $f(x, y) = 0$. L'un et l'autre point sont appelés *points d'extrémum lié*.

2. Extrémums liés et extrémums « absolus »

Etablissons les conditions nécessaires pour que (a, b) soit un point d'extrémum lié d'une fonction différentiable φ sous les hypothèses suivantes concernant f :

1) $f(a, b) = 0$,

2) f est différentiable et f' est continue au point (a, b),

3) le jacobien du système des fonctions $f_1, \ldots, f_m$ est non nul au point (a, b) par rapport à un ensemble au moins de m variables choisies parmi les $n + m$ variables $x_1, \ldots, x_n, y_1, \ldots, y_m$; supposons pour préciser les idées qu'il le soit par rapport aux variables $y_1, \ldots, y_m$:

$$(1) \qquad J(a, b) = \begin{vmatrix} \frac{\partial f_1}{\partial y_1}(a, b) & \ldots & \frac{\partial f_1}{\partial y_m}(a, b) \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \frac{\partial f_m}{\partial y_1}(a, b) & \ldots & \frac{\partial f_m}{\partial y_m}(a, b) \end{vmatrix} \neq 0.$$

Ainsi donc, f vérifie les conditions du théorème 1 du § 1, de sorte qu'il existe des voisinages ouverts U et $V \subset D_f$ des points respectifs $a \in \mathbf{R}^n$ et $(a, b) \in \mathbf{R}^{n+m}$ et une application $g : \mathbf{R}^n \to \mathbf{R}^m$, avec $D_g = U$, tels que l'inter-

section de V et de l'ensemble

$$K = \{(x, y) \in \mathbf{R}^{n+m} \mid f(x, y) = 0\}$$

soit le graphe de l'application g, i.e.

$$V \cap K = \{(x, g(x)) \mid x \in U\}.$$

Ceci étant, g est différentiable et g' est continue au point a. Notons que $V(\subset D_f)$ peut être aussi petit que l'on veut, ce qui découle de la démonstration des théorèmes 9.3.2 à 9.3.4 sur lesquels repose le théorème 1 du § 1.

Posons $\psi(x) = \varphi(x, g(x))$. *Si* (a, b) *est un point de maximum lié* (resp. *de minimum lié*) *de* φ, a *est un point de maximum* (resp. *de minimum*) *de* ψ. En effet, soit par exemple (a, b) un point de maximum lié, de sorte qu'il possède un voisinage ouvert $Q \subset D_\varphi \cap D_f$ tel que

$$\psi(a) = \varphi(a, g(a)) = \varphi(a, b) \geqslant \varphi(x, y) \text{ pour tout } (x, y) \in Q \cap K.$$

Or $(x, g(x)) \in V \cap K$ pour tout $x \in U$ et on peut considérer que V est contenu dans Q. Par suite,

$$\psi(a) \geqslant \varphi(x, g(x)) = \psi(x) \text{ pour tout } x \in U,$$

de sorte que $a = (a_1, \ldots, a_n)$ est un point de maximum de la fonction $\psi(x_1, \ldots, x_n) = \varphi(x_1, \ldots, x_n, g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n))$. La situation avec les minimums lié et « absolu » est analogue.

Mais pour chercher ainsi les points d'extrémum lié, il faut savoir les fonctions $g_1, \ldots, g_m$, i.e. il faut résoudre le système d'équations $f(x, y) = 0$ sans connaître le point (a, b) dans le voisinage duquel on doit le faire. Nous allons exposer la méthode de Lagrange qui détermine les conditions auxquelles le point (a, b) d'extrémum lié doit satisfaire. Ces conditions n'exigent pas la détermination préliminaire de ce point et des fonctions g_k.

3. Méthode des multiplicateurs de Lagrange

On sait que les fonctions g_k existent et sont différentiables. En outre, φ est différentiable par hypothèse. Par suite, ψ est également différentiable en vertu du théorème 7.2.9. Calculons sa différentielle. En vertu du théorème 7.2.9,

$$(2)\quad \frac{\partial \psi}{\partial x_j}(x) = \frac{\partial \varphi}{\partial x_j}(x, g(x)) + \sum_{k=1}^{m} \frac{\partial \varphi}{\partial y_k}(x, g(x)) \frac{\partial g_k}{\partial x_j}(x) \quad (j = 1, \ldots, n),$$

de sorte que

$$(3)\qquad d\psi(x)(u) = \sum_{j=1}^{n} \frac{\partial \psi}{\partial x_j}(x) u_j = \sum_{j=1}^{n} \frac{\partial \varphi}{\partial x_j}(x, g(x)) u_j + \\ + \sum_{k=1}^{m} \frac{\partial \varphi}{\partial y_k}(x, g(x)) \sum_{j=1}^{n} \frac{\partial g_k}{\partial x_j}(x) u_j$$

pour un vecteur arbitraire $u = (u_1, \ldots, u_n) \in \mathbf{R}^n$. Mais $\sum_{j=1}^{n} \frac{\partial g_k}{\partial x_j}(x)u_j = dg_k(x)(u)$. Par conséquent,

$$(4) \quad d\psi(x)(u) = \sum_{j=1}^{n} \frac{\partial \varphi}{\partial x_j}(x, g(x))u_j + \sum_{k=1}^{m} \frac{\partial \varphi}{\partial y_k}(x, g(x))\, dg_k(x)(u).$$

De façon analogue, puisque toutes les fonctions f_i sont différentiables,

$$(5) \quad df_i(x, g(x))(u) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x, g(x))u_j + \sum_{k=1}^{m} \frac{\partial f_i}{\partial y_k}(x, g(x))\, dg_k(x)(u)$$
$$(i = 1, \ldots, m).$$

Mais comme $f(x, g(x)) \equiv 0$, toutes les différentielles $df_i(x, g(x))$ sont nulles. D'autre part, $d\psi$ s'annule au point d'extrémum a. Ainsi donc, le point $(a, b) = (a, g(a))$ doit satisfaire pour tout u aux équations

$$(6) \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x, y)u_j + \sum_{k=1}^{m} \frac{\partial f_i}{\partial y_k}(x, y)\, dg_k(x)(u) = 0 \quad (i = 1, \ldots, m),$$

$$(7) \quad \sum_{j=1}^{n} \frac{\partial \varphi}{\partial x_j}(x, y)u_j + \sum_{k=1}^{m} \frac{\partial \varphi}{\partial y_k}(x, y)\, dg_k(x)(u) = 0.$$

En multipliant les m premières équations par les nombres respectifs $\lambda_1, \ldots, \lambda_m$ (encore indéfinis), et la dernière équation par -1, on obtient après l'addition des résultats obtenus :

$$(8) \quad \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x, y)\lambda_i - \frac{\partial \varphi}{\partial x_j}(x, y) \right) u_j +$$
$$+ \sum_{k=1}^{m} \left(\sum_{i=1}^{m} \frac{\partial f_i}{\partial y_k}(x, y)\lambda_i - \frac{\partial \varphi}{\partial y_k}(x, y) \right) dg_k(x)(u) = 0,$$

équation à laquelle le point recherché $(x, y) = (a, b)$ doit satisfaire pour toutes les valeurs de $\lambda_1, \ldots, \lambda_m$ et $u_1, \ldots, u_n$. Comme à présent $J(a, b) \neq 0$, on peut choisir $\lambda_1, \ldots, \lambda_m$ de telle sorte que les m dernières parenthèses dans (8) s'annulent (et les fonctions g_k inconnues soient donc exclues !), i.e. les équations

$$\sum_{i=1}^{m} \frac{\partial f_i}{\partial y_k}(a, b)\lambda_i = \frac{\partial \varphi}{\partial y_k}(a, b) \qquad (k = 1, \ldots, m)$$

soient satisfaites. En effet, c'est un système de m équations du premier degré à m inconnues $\lambda_1, \ldots, \lambda_m$ dont le déterminant est $J(a, b) \neq 0$. Pour un tel choix des nombres il résulte de l'équation (8) que

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(a, b)\lambda_i - \frac{\partial \varphi}{\partial x_j}(a, b) \right) u_j = 0$$

quel que soit le vecteur u. En posant $u = e^j = (\delta_1^j, \ldots, \delta_n^j)$ (où δ_i^j sont les symboles de Kronecker), on obtient

$$\sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(a, b)\lambda_i - \frac{\partial \varphi}{\partial x_j}(a, b) = 0 \qquad \text{pour tout entier } j \in [1, n].$$

Ainsi donc, les coordonnées du point (a, b) d'extrémum lié doivent, pour les valeurs données des nombres λ, annuler les $n + m$ parenthèses dans l'équation (8). Mais ces parenthèses sont les dérivées partielles de la fonction

$$F = \lambda_1 f_1 + \ldots + \lambda_m f_m - \varphi \tag{9}$$

par rapport à tous les arguments $x_1, \ldots, x_n, y_1, \ldots, y_m$, et leur annulation est une condition nécessaire d'existence du point d'extrémum de cette fonction. Le résultat obtenu justifie la règle suivante.

RÈGLE DE LAGRANGE. *Pour trouver les points éventuels d'extrémum d'une fonction* $\varphi(x, y)$ (*où* $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$) *soumise à des contraintes égalité* $f_1(x, y) = 0, \ldots, f_m(x, y) = 0$, *il faut composer la fonction lagrangienne* (9) *avec des coefficients réels* (*encore indéfinis*) $\lambda_1, \ldots, \lambda_m$ *et écrire les conditions nécessaires de son extrémum* :

$$\frac{\partial F}{\partial x_j}(x, y) \equiv \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x, y)\lambda_i - \frac{\partial \varphi}{\partial x_j}(x, y) = 0 \quad (j = 1, \ldots, n) \tag{10}$$

et

$$\frac{\partial F}{\partial y_k}(x, y) \equiv \sum_{i=1}^{m} \frac{\partial f_i}{\partial y_k}(x, y)\lambda_i - \frac{\partial \varphi}{\partial y_k}(x, y) = 0 \quad (k = 1, \ldots, m). \tag{11}$$

Avec les m contraintes, cela donne $n + 2m$ *équations auxquelles doivent satisfaire m multiplicateurs* $\lambda_1, \ldots, \lambda_m$ *et* $n + m$ *coordonnées* $x_1, \ldots, x_n$, $y_1, \ldots, y_m$ *du point d'extrémum lié cherché.*

Vu que les conditions (10) et (11) pour les dérivées par rapport à x_j et y_k ont la même forme, on peut ne pas distinguer ces arguments des fonctions φ et f_i, i.e. choisir, au lieu de $y_1, \ldots, y_m$, les m arguments quelcon-

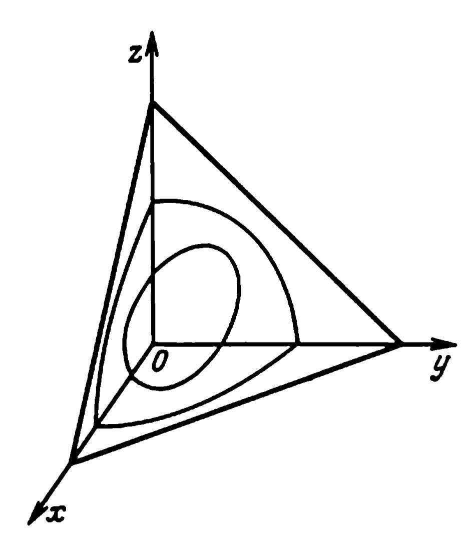


Fig. 31

ques parmi $x_1, \ldots, x_n, y_1, \ldots, y_m$ (pourvu que le jacobien du système des fonctions f_i par rapport à ces arguments ne s'annule pas).

Proposons-nous par exemple de trouver les valeurs maximale et minimale que la fonction

$$\varphi(x, y, z) = xyz$$

présente sur une circonférence K qui est l'intersection du plan

$$x + y + z = 5 \tag{12}$$

et de la sphère

$$x^2 + y^2 + z^2 = 9 \tag{13}$$

(fig. 31). Le plan (12) coupe en effet la sphère (13) puisque la distance de l'origine des coordonnées à ce plan $\left(\text{égale *) à } \frac{5}{\sqrt{3}}\right)$ est inférieure au rayon de la sphère. La circonférence K est fermée (en tant qu'intersection de deux ensembles fermés) et bornée (parce que contenue dans la sphère (13)), si bien que K est un compact. C'est pourquoi, φ présente sur K les valeurs maximale et minimale. Les points où ces valeurs sont atteintes, sont des points d'extrémum lié de la fonction φ sous les contraintes (12) et (13). L'application $f : \mathbf{R}^3 \to \mathbf{R}^2$ de composantes

$$f_1(x, y, z) = x + y + z - 5, \quad f_2(x, y, z) = x^2 + y^2 + z^2 - 9$$

satisfait dans ce cas aux conditions 1) à 3) formulées au début du n° 2. En effet : 1) $f(P) = 0$ pour tout point $P \in K$ en vertu de (12) et (13) ; 2) f est continûment différentiable ainsi que f_1 et f_2 ; 3) la matrice jacobienne de l'application f est

$$\begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}$$

*) On sait du cours de Géométrie analytique que la distance de l'origine des coordonnées au plan $Ax + By + Cz + D = 0$ est égale à $\frac{D}{\sqrt{A^2 + B^2 + C^2}}$ (ce qu'on peut établir en recherchant le minimum de la fonction $x^2 + y^2 + z^2$ sous la contrainte $Ax + By + Cz + D = 0$).

et, par conséquent, les jacobiens du système des fonctions f_1, f_2 par rapport aux couples de variables (x, y), (x, z) et (y, z) sont respectivement les déterminants

$$\begin{vmatrix} 1 & 1 \\ 2x & 2y \end{vmatrix} = 2(y - x), \quad \begin{vmatrix} 1 & 1 \\ 2x & 2z \end{vmatrix} = 2(z - x), \quad \begin{vmatrix} 1 & 1 \\ 2y & 2z \end{vmatrix} = 2(z - y)$$

qui ne s'annulent simultanément que sur la droite $x = y = z$; or aucun point de cette droite ne vérifie les équations (12) et (13) puisque les égalités $3x = 5$ et $3x^2 = 9$ sont incompatibles. Ainsi donc, pour trouver les points d'extrémum lié, on peut appliquer la règle de Lagrange. Ecrivons la fonction lagrangienne

$$\lambda(x + y + z - 5) + \mu(x^2 + y^2 + z^2 - 9) - xyz$$

et annulons ses dérivées partielles par rapport à x, y, z. On obtient les équations

(14′) $$\lambda + 2\mu x - yz = 0,$$

(14″) $$\lambda + 2\mu y - xz = 0,$$

(14‴) $$\lambda + 2\mu z - xy = 0$$

qui, jointes aux équations (12) et (13), doivent être vérifiées par les coefficients λ, μ et par les coordonnées x, y, z des points d'extrémum lié. En additionnant les équations (14′), (14″) et (14‴) et en prenant en considération qu'en vertu de (12) et de (13)

$$yz + xz + xy = \frac{1}{2}\left[(x + y + z)^2 - (x^2 + y^2 + z^2)\right] = 8,$$

on obtient

(15) $$3\lambda + 10\mu = 8.$$

D'autre part, en retranchant (14″) de (14′), (14‴) de (14″), et (14′) de (14‴), on a

(16) $$\begin{cases} (2\mu + z)(x - y) = 0, \\ (2\mu + x)(y - z) = 0, \\ (2\mu + y)(z - x) = 0. \end{cases}$$

Comme les égalités $x = y = z$ et le système (12)-(13) sont incompatibles, les deuxièmes parenthèses dans (16) ne sont pas simultanément non nulles (car, autrement, les trois premières seraient nulles, d'où il résulterait que $x = y = z$) et de plus une seule des deuxièmes parenthèses est nulle (car l'annulation de deux quelconques d'entre elles entraînerait de nouveau $x = y = z$). Soit, par exemple, $x - y = 0$. Comme dans ce cas $y - z \neq 0$ et $z - x \neq 0$, on obtient $2\mu + x = 2\mu + y$, d'où $2\mu = -x = -y$. Il résulte alors de l'équation (12) que $z = 5 - 2x$ et l'égalité (14′) prend la forme $\lambda - x^2 - x(5 - 2x) = 0$, d'où $\lambda = 5x - x^2$. La substitution des valeurs trouvées de λ et de 2μ dans la formule (15) donne l'équation

$$3x^2 - 10x + 8 = 0,$$

d'où on obtient deux valeurs de x : $x_1 = 2$ et $x_2 = \frac{4}{3}$ et, par conséquent, deux points éventuels d'extrémum lié : $(2, 2, 1)$ et $\left(\frac{4}{3}, \frac{4}{3}, \frac{7}{3}\right)$. On obtient de la même façon pour $y - z = 0$ et pour $z - x = 0$ les points $(2, 1, 2)$ et $\left(\frac{4}{3}, \frac{7}{3}, \frac{4}{3}\right)$ et, respectivement, $(1, 2, 2)$ et

$\left(\frac{7}{3}, \frac{4}{3}, \frac{4}{3}\right)$. Puisque xyz ne dépend pas de l'ordre de ses facteurs, on peut conclure que sa valeur maximale sur K est égale à $\frac{4}{3} \cdot \frac{4}{3} \cdot \frac{7}{3}$, i.e. à $4 \cdot \frac{28}{27}$, et la valeur minimale est égale à $2 \cdot 2 \cdot 1$, i.e. à 4, et chacune d'elles est atteinte en trois points : la première aux points $\left(\frac{4}{3}, \frac{4}{3}, \frac{7}{3}\right)$, $\left(\frac{4}{3}, \frac{7}{3}, \frac{4}{3}\right)$ et $\left(\frac{7}{3}, \frac{4}{3}, \frac{4}{3}\right)$ et la deuxième aux points (2, 2, 1), (2, 1, 2) et (1, 2, 2).

Troisième partie

CALCUL INTÉGRAL DES FONCTIONS DE DEUX ET DE TROIS VARIABLES

CHAPITRE 11

INTÉGRALE DOUBLE

§ 1. Renseignements supplémentaires sur les figures quarrables

1. Notion de figure quarrable et propriétés élémentaires

Etant donné le plan de coordonnées Π, considérons pour tout entier $n \geqslant 0$ une subdivision de Π par les droites

$$x = \frac{k}{2^n}, \; y = \frac{l}{2^n} \qquad (k, l \in \mathbf{Z})$$

en « carrés » congruents $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \times \left[\frac{l}{2^n}, \frac{l+1}{2^n}\right]$ de rang n. Soit Φ une figure (i.e. un ensemble borné) dans Π. Il existe un ensemble fini (peut-être vide) de carrés de rang n contenus dans Φ et un ensemble fini (vide si Φ est vide) de carrés de rang n qui se coupent avec Φ. Soient $A_n(\Phi)$ la réunion des premiers carrés, et $B_n(\Phi)$ celle des deuxièmes, et soient $\alpha_n(\Phi)$ le nombre des premiers carrés, et $\beta_n(\Phi)$ celui des deuxièmes. Etant donné

$$p_n(\Phi) = \frac{\alpha_n(\Phi)}{4^n}, \qquad q_n(\Phi) = \frac{\beta_n(\Phi)}{4^n},$$

on démontre que $(p_n(\Phi))$ est une suite strictement croissante et $(q_n(\Phi))$, une suite strictement décroissante, et que $p_m(\Phi) \leqslant q_n(\Phi)$ pour tous $m, n \geqslant 0$. Il en résulte que :

1) la borne supérieure de la suite $(p_m(\Phi))$ est finie ; on l'appelle *aire intérieure* de la figure Φ et on la note $\underline{\text{aire}}\,\Phi$;

2) la borne inférieure de la suite $(q_n(\Phi))$ est finie ; on l'appelle *aire extérieure* de la figure Φ et on la note $\overline{\text{aire}}\,\Phi$;

3) $\underline{\text{aire}}\,\Phi \leqslant \overline{\text{aire}}\,\Phi$.

Si $\underline{\text{aire}}\,\Phi = \overline{\text{aire}}\,\Phi$, on dit que Φ est une *figure quarrable* et que la valeur commune de ses aires intérieure et extérieure est son *aire* qu'on note aire Φ. Ceci étant, aire $\Phi = \lim p_n(\Phi) = \lim q_n(\Phi)$ car, en général, $\lim p_n(\Phi) = \underline{\text{aire}}\,\Phi$ et $\lim q_n(\Phi) = \overline{\text{aire}}\,\Phi$.

Si les figures Φ et Ψ sont quarrables et $\Phi \subset \Psi$, aire $\Phi \leqslant$ aire Ψ (monotonie de l'aire).

Si les figures Φ et Ψ sont quarrables et n'ont pas de points intérieurs communs, leur réunion $\Phi \cup \Psi$ est quarrable et aire $(\Phi \cup \Psi)$ = aire Φ + + aire Ψ (additivité de l'aire).

Pour qu'une figure Φ soit quarrable, il est nécessaire et suffisant que pour tout $\varepsilon > 0$ il existe des figures quarrables A et B telles que $A \subset \Phi \subset B$ et aire B − aire $A < \varepsilon$.

Tous les polygones sont quarrables et leurs aires sont calculées d'après les formules usuelles. En particulier, si Φ est un carré de rang n, $q_n(\Phi)$ = aire $B_n(\Phi)$.

Enfin, la quarrabilité et l'aire de la figure restent invariantes dans les transformations orthogonales du plan Π et, par conséquent, ne dépendent pas du choix du système de coordonnées (avec la même échelle de mesure des longueurs).

2. Figures d'aire nulle

La figure Φ *est quarrable et* aire $\Phi = 0$ *si et seulement si* $\overline{\text{aire}}\ \Phi = 0$ (car on a alors $\underline{\text{aire}}\ \Phi = 0$, de sorte que $\underline{\text{aire}}\ \Phi = \overline{\text{aire}}\ \Phi$), *i.e. si* $q_n(\Phi) \to 0$. Il est évident que *toute partie d'une figure d'aire nulle est quarrable et son aire est nulle.*

LEMME 1. *Quelles que soient les figures* $\Phi_1, \ldots, \Phi_m$ *du plan* Π,

$$\overline{\text{aire}} \left(\bigcup_{i=1}^{m} \Phi_i \right) \leqslant \sum_{i=1}^{m} \overline{\text{aire}}\ \Phi_i. \tag{1}$$

Démonstration. Soit $\Phi = \bigcup_{i=1}^{m} \Phi_i$. Vu que tout carré de rang n qui a des points communs avec Φ coupe un Φ_i au moins, $\beta_n(\Phi) \leqslant \sum_{i=1}^{m} \beta_n(\Phi_i)$, d'où $q_n(\Phi) \leqslant \sum_{i=1}^{m} q_n(\Phi_i)$. Si $n \to \infty$, on obtient à la limite l'inégalité (1).

Il résulte du lemme 1 que *la réunion de toute famille finie de figures d'aire nulle est une figure d'aire nulle.*

THÉORÈME 1. *L'ensemble L des points de l'arc de Jordan rectifié* $\mathscr{L}$ *est une figure d'aire nulle.*

Démonstration. Supposons que $\mathscr{L}$ soit donné par les équations paramétriques

$$x = \varphi(t), \qquad y = \psi(t) \qquad (a \leqslant t \leqslant b)$$

où φ et ψ sont des fonctions continues sur le segment $[a, b]$. Comme $\mathscr{L}$ est rectifiable, on démontre que ψ est une fonction à variation bornée, de sorte que $\psi = \psi_1 - \psi_2$ où ψ_1 et ψ_2 sont des fonctions strictement croissan-

tes. Soit $\varepsilon > 0$ et

(2) $$\eta = \frac{\varepsilon}{[\psi_1(b) - \psi_1(a)] + [\psi_2(b) - \psi_2(a)] + \varepsilon}.$$

Le corollaire du théorème 3.2.14 dit que φ est uniformément continue sur le segment $[a, b]$, si bien qu'il existe un $m \in \mathbf{N}$ tel que

(3) $$\left(t', t'' \in [a, b] \text{ et } |t' - t''| \leqslant \frac{b-a}{m}\right) \Rightarrow |\varphi(t') - \varphi(t'')| \leqslant \eta.$$

Les points $t_i = a + \frac{b-a}{m} i$ $(i = 0, 1, \ldots, m)$ divisent $[a, b]$ en segments $\Delta_i = [t_i, t_{i+1}]$ $(i = 0, 1, \ldots, m-1)$ de longueur $\frac{b-a}{m}$. Les fonctions φ et ψ présentent, sur chaque segment Δ_i, les valeurs minimale et maximale. Supposons que ce soient x_i' et x_i'' pour φ, et y_i' et y_i'' pour ψ. Posons

$$\mathscr{P}_i = [x_i', x_i''] \times [y_i', y_i''].$$

Comme $x_i' = \varphi(t_i')$ et $x_i'' = \varphi(t_i'')$, où $t_i', t_i'' \in \Delta_i$, et donc $|t_i' - t_i''| \leqslant \frac{b-a}{m}$, on obtient en vertu de (3) que $x_i'' - x_i' \leqslant \eta$. D'autre part, comme $y_i' = \psi(\tau_i')$ et $y_i'' = \psi(\tau_i'')$ où $\tau_i', \tau_i'' \in \Delta_i$, on a d'après les propriétés de la variation d'une fonction :

$$y_i'' - y_i' \leqslant \bigvee_{\min(\tau_i', \tau_i'')}^{\max(\tau_i', \tau_i'')} (\psi_1) \leqslant \bigvee_{t_i}^{t_{i+1}} (\psi) \leqslant \bigvee_{t_i}^{t_{i+1}} (\psi_1) + \bigvee_{t_i}^{t_{i+1}} (\psi_2) =$$
$$= [\psi_1(t_{i+1}) - \psi_1(t_i)] + [\psi_2(t_{i+1}) - \psi_2(t_i)].$$

Par suite,

(4) $$\text{aire}\, \mathscr{P}_i = (x_i'' - x_i')(y_i'' - y_i') \leqslant$$
$$\leqslant \eta\{[\psi_1(t_{i+1}) - \psi_1(t_i)] + [\psi_2(t_{i+1}) - \psi_2(t_i)]\}.$$

Posons $\mathscr{P} = \bigcup_{i=0}^{m-1} \mathscr{P}_i$. Soit $(x, y) \in L$, i.e. il existe un $t \in [a, b]$ tel que $x = \varphi(t)$, $y = \psi(t)$. Comme $t \in \Delta_i$ pour un i, on a $(x, y) \in \mathscr{P}_i$ puisque $x_i' \leqslant \varphi(t) \leqslant x_i''$ et $y_i' \leqslant \psi(t) \leqslant y_i''$. Ainsi donc, $L \subset \mathscr{P}$, d'où $\overline{\text{aire}}\, L \leqslant \overline{\text{aire}}\, \mathscr{P}$. Or $\overline{\text{aire}}\, \mathscr{P} \leqslant \sum_{i=0}^{m-1} \overline{\text{aire}}\, (\mathscr{P}_i)$ suivant le lemme 1. Par conséquent, on obtient en vertu de (1) que

$$(0 \leqslant)\ \overline{\text{aire}}\, L \leqslant \sum_{i=0}^{m-1} \eta\{[\psi_1(t_{i+1}) - \psi_1(t_i)] + [\psi_2(t_{i+1}) - \psi(t_i)]\} =$$
$$= \eta\{[\psi_1(b) - \psi_1(a)] + [\psi_2(b) - \psi_2(a)]\} < \varepsilon$$

et donc $\overline{\text{aire}}\ L = 0$ puisque ε est arbitraire, i.e. L est une figure d'aire nulle.

3. Quarrabilité et frontière *)

THÉORÈME 2. *La figure Φ est quarrable si et seulement si sa frontière est quarrable et l'aire de la frontière est nulle.*

Démonstration. Rappelons que $A_n(\Phi)$ est la réunion des carrés de rang n contenus dans $]\Phi[$, $\alpha_n(\Phi)$ est le nombre de ces carrés, $B_n(\Phi)$ est la réunion des carrés de rang n ayant des points communs avec Φ, et $\beta_n(\Phi)$ le nombre de ces carrés. Notons $C_n(\Phi)$ et $\gamma_n(\Phi)$ respectivement la réunion et le nombre de carrés de rang n ayant des points communs avec Φ, mais qui ne sont pas contenus dans $]\Phi[$. Alors

$$A_n(\Phi) \cup C_n(\Phi) = B_n(\Phi) \tag{5}$$

et, puisque $\alpha_n(\Phi) + \gamma_n(\Phi) = \beta_n(\Phi)$,

$$\text{aire}\ A_n(\Phi) + \text{aire}\ C_n(\Phi) = \text{aire}\ B_n(\Phi). \tag{6}$$

Comme $B_n(\Phi)$ est fermé (en tant que réunion d'une famille finie de carrés) et $\Phi \subset B_n(\Phi)$, on obtient $[\Phi] \subset B_n(\Phi)$ et, à plus forte raison, $\text{Fr}\ \Phi \subset B_n(\Phi)$. D'autre part, $\text{Fr}\ \Phi \cap A_n(\Phi) = \varnothing$. Ainsi donc, $\text{Fr}\ \Phi \subset B_n(\Phi) \setminus A_n(\Phi)$. Or $B_n(\Phi) \setminus A_n(\Phi) \subset C_n(\Phi)$ en vertu de (5). Par conséquent,

$$\text{Fr}\ \Phi \subset C_n(\Phi). \tag{7}$$

D'autre part,

$$C_n(\Phi) \subset B_n(\text{Fr}\ \Phi). \tag{8}$$

En effet, soit Q un des carrés de rang n qui forment $C_n(\Phi)$. Il existe alors dans Q un point $a \in \Phi$ et un point $b \notin]\Phi[$. Si $b \in \Phi$, on a $b \in \text{Fr}\ \Phi$, de sorte que Q a des points communs avec $\text{Fr}\ \Phi$, i.e. il est contenu dans $B_n(\text{Fr}\ \Phi)$. Si, par contre, $b \notin \Phi$, on a $a \neq b$ et, d'après le théorème 3.3.2, le segment $[a, b]$ contient un point $c \in \text{Fr}\ \Phi$; mais comme Q est convexe, $[a, b] \subset Q$ et, par conséquent, $c \in Q$, de sorte que Q a de nouveau une intersection non vide avec $\text{Fr}\ \Phi$. Il résulte de (7), (6) et (8) que

$$\overline{\text{aire}}\ \text{Fr}\ \Phi \leqslant \text{aire}\ C_n(\Phi) = \text{aire}\ B_n(\Phi) - \text{aire}\ A_n(\Phi) \leqslant \text{aire}\ B_n(\text{Fr}\ \Phi).$$

En faisant tendre n vers l'infini, on obtient à la limite

$$\overline{\text{aire}}\ \text{Fr}\ \Phi \leqslant \lim\ (\text{aire}\ B_n(\Phi)) - \lim\ (\text{aire}\ A_n(\Phi)) \leqslant \lim\ (\text{aire}\ B_n(\text{Fr}\ \Phi)),$$

i.e. $\overline{\text{aire}}\ \text{Fr}\ \Phi \leqslant \overline{\text{aire}}\ \Phi - \underline{\text{aire}}\ \Phi \leqslant \overline{\text{aire}}\ \text{Fr}\ \Phi$. Par suite,

$$\overline{\text{aire}}\ \text{Fr}\ \Phi = \overline{\text{aire}}\ \Phi - \underline{\text{aire}}\ \Phi.$$

---

*) La notion et les propriétés de la frontière d'un ensemble sont envisagées au § 3 du chapitre 3.

Il en résulte immédiatement que Φ est quarrable, i.e. $\overline{\text{aire}}\ \Phi = \underline{\text{aire}}\ \Phi$, si et seulement si $\overline{\text{aire}}\ \text{Fr}\ \Phi = 0$, i.e. $\text{Fr}\ \Phi$ est quarrable et son aire est nulle.

COROLLAIRE. *La figure limitée par un arc de Jordan rectifié ou par une famille finie de tels arcs est quarrable.*

En effet, en vertu du théorème 1, sa frontière est une figure d'aire nulle.

En particulier, *la figure limitée par un ou plusieurs arcs de classe C^1 est quarrable.* En analyse mathématique on ne rencontre habituellement que des figures de ce type.

THÉORÈME 3. *La réunion, l'intersection et la différence de deux figures quarrables sont des figures quarrables.*

Démonstration. Supposons que Φ soit $\Phi_1 \cup \Phi_2$, $\Phi_1 \cap \Phi_2$ ou $\Phi_1 \setminus \Phi_2$. D'après le théorème 3.3.1, dans les trois cas on a $\text{Fr}\ \Phi \subset \text{Fr}\ \Phi_1 \cup \text{Fr}\ \Phi_2$. Si Φ_1 et Φ_2 sont quarrables, le théorème 2 dit que $\text{Fr}\ \Phi_1$ et $\text{Fr}\ \Phi_2$ sont des figures d'aire nulle. Il s'ensuit que $\text{Fr}\ \Phi_1 \cup \text{Fr}\ \Phi_2$ et $\text{Fr}\ \Phi$ sont aussi des figures d'aire nulle, donc, de nouveau suivant le théorème 2, Φ est quarrable.

REMARQUE. L'assertion du théorème 3 s'étend par récurrence à la réunion et à l'intersection de toute famille finie de figures quarrables. En particulier, si les figures $\Phi_1, \ldots, \Phi_m$ sont quarrables, la figure $\Phi = \bigcup_{i=1}^{m} \Phi_i$ est aussi quarrable et $\text{aire}\ \Phi \leqslant \sum_{i=1}^{m} \text{aire}\ \Phi_i$. En effet, en vertu du lemme 1,

$$\text{aire}\ \Phi = \overline{\text{aire}}\ \Phi \leqslant \sum_{i=1}^{m} \overline{\text{aire}}\ \Phi_i = \sum_{i=1}^{m} \text{aire}\ \Phi_i.$$

LEMME 2. $[\Phi] = \Phi \cup \text{Fr}\ \Phi$.

Démonstration. Vu que $\Phi \subset [\Phi]$ et $\text{Fr}\ \Phi \subset [\Phi]$, on a $\Phi \cup \text{Fr}\ \Phi \subset$ $\subset [\Phi]$. D'autre part, comme $[\Phi] \setminus \Phi = [\Phi] \cap \mathbf{C}\Phi \subset [\Phi] \cap [\mathbf{C}\Phi] = \text{Fr}\ \Phi$, on obtient $[\Phi] = \Phi \cup ([\Phi] \setminus \Phi) \subset \Phi \cup \text{Fr}\ \Phi$.

THÉORÈME 4. *L'adhérence $[\Phi]$ d'une figure quarrable Φ est quarrable et possède la même aire que Φ.*

Démonstration. Etant donné que Φ est quarrable, $\text{Fr}\ \Phi$ est aussi quarrable selon le théorème 2. Or $[\Phi] = \Phi \cup \text{Fr}\ \Phi$ en vertu du lemme 2. Par conséquent, d'après le théorème 3, $[\Phi]$ est également quarrable. Puisque $\text{aire Fr}\ \Phi = 0$ suivant le théorème 2, on a, en vertu des lemmes 1 et 2,

$$\text{aire}\ \Phi \leqslant \text{aire}\ [\Phi] \leqslant \text{aire}\ \Phi + \text{aire Fr}\ \Phi = \text{aire}\ \Phi,$$

i.e. $\text{aire}\ [\Phi] = \text{aire}\ \Phi$.

THÉORÈME 5. *L'adhérence d'une figure quarrable est un compact quarrable.*

Démonstration. En tant qu'adhérence, cet ensemble est fermé, et en tant qu'adhérence d'un ensemble borné, il est borné. Or tout ensemble

fermé et borné dans le plan est un compact. Il est quarrable d'après le théorème 4.

A noter que dans le plan il existe encore des compacts non quarrables (dont la construction est assez compliquée).

§ 2. Corps solides cubables

1. Notion de corps solide cubable

L'espace de coordonnées est subdivisé pour tout entier $n \geqslant 0$ par les plans

$$x = \frac{k}{2^n}, \qquad y = \frac{l}{2^n}, \qquad z = \frac{m}{2^n} \qquad (k, l, m \in \mathbf{Z})$$

en « cubes » congruents de rang n :

$$\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \times \left[\frac{l}{2^n}, \frac{l+1}{2^n}\right] \times \left[\frac{m}{2^n}, \frac{m+1}{2^n}\right].$$

Appelons *corps solide* tout ensemble borné de points de l'espace. Soit $\mathscr{T}$ un corps solide arbitraire. Il existe un ensemble fini (peut-être vide) de cubes de rang n contenus à l'intérieur de $\mathscr{T}$, et un ensemble fini (vide si $\mathscr{T}$ est vide) de cubes de rang n qui se coupent avec $\mathscr{T}$. Soient $A_n(\mathscr{T})$ la réunion des premiers cubes et $B_n(\mathscr{T})$, celle des deuxièmes, et soient $\alpha_n(\mathscr{T})$ le nombre des premiers cubes et $\beta_n(\mathscr{T})$ celui des deuxièmes. Etant donné

$$p_n(\mathscr{T}) = \frac{\alpha_n(\mathscr{T})}{8^n}, \qquad q_n(\mathscr{T}) = \frac{\beta_n(\mathscr{T})}{8^n},$$

on établit comme dans le cas des figures planes que $(p_n(\mathscr{T}))$ est une suite strictement croissante et que $(q_n(\mathscr{T}))$ est une suite strictement décroissante. On a de plus $p_m(\mathscr{T}) \leqslant q_n(\mathscr{T})$ pour tous $m, n \geqslant 0$, d'où il résulte que :

1) la borne supérieure de la suite $(p_n(\mathscr{T}))$ est finie ; on l'appelle *volume intérieur* du corps $\mathscr{T}$ et on la note $\underline{\mathrm{vol}}\, \mathscr{T}$;

2) la borne inférieure de la suite $(q_n(\mathscr{T}))$ est finie ; on l'appelle *volume extérieur* du corps $\mathscr{T}$ et on la note $\overline{\mathrm{vol}}\, \mathscr{T}$;

3) $\underline{\mathrm{vol}}\, \mathscr{T} \leqslant \overline{\mathrm{vol}}\, \mathscr{T}$.

Si $\underline{\mathrm{vol}}\, \mathscr{T} = \overline{\mathrm{vol}}\, \mathscr{T}$, le corps $\mathscr{T}$ est dit *cubable*, la valeur commune de ses volumes extérieur et intérieur s'appelle *volume* tout court ; il sera noté $\mathrm{vol}\, \mathscr{T}$.

2. Propriétés fondamentales des corps solides cubables

La théorie des corps cubables est élaborée de façon analogue que celle des figures quarrables ; certaines modifications sont dues au fait que nous possédons maintenant la notion de frontière d'un ensemble. Voici les étapes essentielles de construction de cette théorie.

Tout comme dans le cas des figures planes, on démontre que le volume est *additif* et *monotone* et que le parallélépipède rectangle

(1) $$[a_1, a_2] \times [b_1, b_2] \times [c_1, c_2]$$

est cubable et son volume est $(a_2 - a_1)(b_2 - b_1)(c_2 - c_1)$. Par analogie au cas plan, il en résulte que la cubabilité et le volume sont invariants par translation.

Analogiquement au cas des figures planes (voir théorème 2 du § 1) on démontre deux critères suivants de cubabilité :

1) le corps $\mathcal{F}$ est cubable si et seulement si pour tout $\varepsilon > 0$ il existe des corps cubables $\mathcal{A}$ et $\mathcal{B}$ tels que $\mathcal{A} \subset \mathcal{F} \subset \mathcal{B}$ et vol $\mathcal{B}$ − vol $\mathcal{A} < \varepsilon$;

2) le corps $\mathcal{F}$ est cubable si et seulement si sa frontière est cubable et a un volume nul.

Tout comme le théorème 3 du § 1, il résulte du deuxième critère que la réunion, l'intersection et la différence de deux corps cubables sont des corps cubables. Le même critère permet de démontrer la cubabilité des polyèdres un peu autrement que la quarrabilité des polygones.

LEMME. *Si le corps $\mathcal{F}$ est cubable et possède un volume nul, il en est de même de son image $U\mathcal{F}$ par toute transformation orthogonale U de l'espace.*

Démonstration. Tout cube de rang n ayant des points communs avec $U\mathcal{F}$ se coupe avec $UB_n(\mathcal{F})$, i.e. avec UC où C est un cube de rang n ayant des points communs avec $\mathcal{F}$. Or UC se coupe avec 27 cubes de rang n au plus. En effet, UC est un cube congruent à C ; soit

$$C' = \left[\frac{k'}{2^n}, \frac{k'+1}{2^n}\right] \times \left[\frac{l'}{2^n}, \frac{l'+1}{2^n}\right] \times \left[\frac{m'}{2^n}, \frac{m'+1}{2^n}\right]$$

un cube de rang n qui contient le centre w_0 de UC. La distance de chaque point de UC à w_0 est au plus égale à la moitié de la longueur $\sqrt{3}/2^n$ de sa diagonale et, par conséquent, cette distance est inférieure à $1/2^n$. C'est pourquoi, aucun cube C'' de rang n ayant des points communs avec UC ne peut se trouver en dehors d'un cube « triplé » (dans toutes les directions)

$$\left[\frac{k'-1}{2^n}, \frac{k'+2}{2^n}\right] \times \left[\frac{l'-1}{2^n}, \frac{l'+2}{2^n}\right] \times \left[\frac{m'-1}{2^n}, \frac{m'+2}{2^n}\right]$$

car il n'y existe pas de points dont la distance aux points de C' (en particulier, à w_0) soit inférieure à $1/2^n$. Mais ce cube « triplé » est composé évidemment de 27 cubes de rang n avec lesquels peuvent se couper seulement de tels C'' (et, par conséquent, le cube UC).

Il en résulte que $\beta_n(U\mathcal{F}) \leqslant 27\beta_n(\mathcal{F})$ et, par conséquent, $q_n(U\mathcal{F}) \leqslant 27q_n(\mathcal{F})$. Mais vol $\mathcal{F} = 0$ implique $q_n(\mathcal{F}) \to 0$. Donc, $q_n(U\mathcal{F}) \to 0$, i.e. le corps $U\mathcal{F}$ est cubable et vol $U\mathcal{F} = 0$.

COROLLAIRE. *Toute figure plane Φ dans l'espace est cubable et possède un volume nul.*

En effet, Φ est l'image par transformation orthogonale dans l'espace d'une figure congruente (quelconque) Φ_1 située dans le plan xOy. A son tour, Φ_1 est contenue dans un rectangle $[a_1, a_2] \times [b_1, b_2] \subset xOy$ et, par conséquent, dans un parallélépipède rectangle (1) de hauteur $c_2 - c_1$ aussi petite que l'on veut, i.e. de volume aussi petit que l'on veut. Donc, Φ_1 est cubable et possède un volume nul, ce qui implique d'après le lemme qu'il en est de même pour Φ.

Il en résulte que *tout polyèdre* (*et, en particulier, tout cube*) *est cubable* : sa frontière est la réunion d'une famille finie de figures planes, et, par conséquent, elle est cubable et son volume est nul. En se fondant sur ce fait, on établit, par analogie à la démonstration de ce que la quarrabilité et l'aire sont invariantes par une symétrie orthogonale dans le plan par rapport à une droite, que la cubabilité et le volume sont invariants par une symétrie orthogonale dans l'espace par rapport à un plan. Et comme chaque transformation orthogonale de l'espace est la composée de plusieurs symétries de ce type, il en résulte que la cubabilité et le volume sont invariants par toute transformation orthogonale, c.-à-d. qu'ils ne dépendent pas du choix du système de coordonnées (échelle de mesure de la longueur étant la même).

3. Relation entre les aires et les volumes

Soient Φ une figure arbitraire dans le plan xOy et $[m, M]$ un segment arbitraire de l'axe Oz. On dira que

$$\Phi \times [m, M] = \{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in \Phi, \quad m \leqslant z \leqslant M\}$$

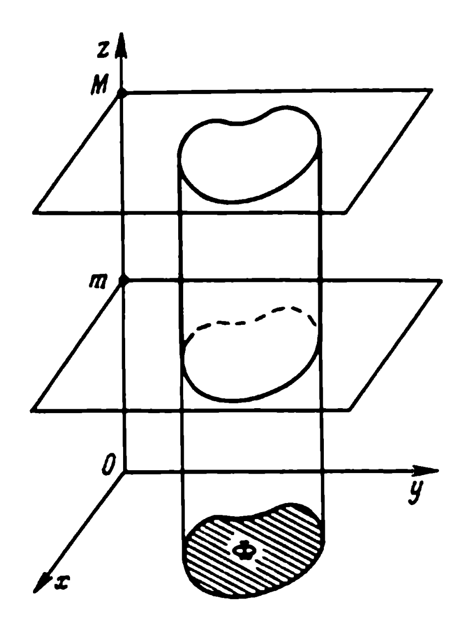


Fig. 32

est un *corps cylindrique* (fig. 32). Le théorème suivant établit une relation entre les notions d'aire et de volume.

THÉORÈME 1. *Si la figure* Φ *est quarrable, le corps cylindrique* $\mathcal{T} = \Phi \times [m, M]$ *est cubable et* $\operatorname{vol} \mathcal{T} = (M - m)$ aire Φ.

Démonstration. Etant donné que $\mathcal{T} = \Phi \times [m, M]$, on a

$$A_n(\Phi) \times [m, M] \subset \mathcal{T} \subset B_n(\Phi) \times [m, M]$$

où $A_n(\Phi)$ et $B_n(\Phi)$ sont respectivement les ensembles des carrés de rang n du plan xOy contenus à l'intérieur de Φ et ayant des points communs avec Φ. Les corps cylindriques $A_n(\Phi) \times [m, M]$ et $B_n(\Phi) \times [m, M]$ contiennent un nombre fini de parallélépipèdes rectangles qui n'ont pas de points intérieurs communs, de sorte qu'ils sont cubables et

$$\operatorname{vol}(A_n(\Phi) \times [m, M]) = (M - m) \text{ aire } A_n(\Phi),$$
$$\operatorname{vol}(B_n(\Phi) \times [m, M]) = (M - m) \text{ aire } B_n(\Phi).$$

La figure Φ étant quarrable, les deux volumes tendent vers $(M - m)$ aire Φ. Par conséquent, le corps $\mathcal{T}$ est cubable et $(M - m)$ aire Φ est son volume.

Il est évident que le théorème 1 est vrai dans les cas où Φ est une figure quarrable dans le plan yOz ou xOz et $[m, M]$ est un segment de l'axe Ox ou Oy respectivement.

§ 3. Notion d'intégrale double

L'intégrale définie est une intégrale sur un segment. L'intégrale double est une intégrale sur un compact quarrable. Tous les compacts quarrables envisagés plus loin sont non vides ; certains d'entre eux peuvent avoir l'aire non nulle.

1. Partition des compacts quarrables

On appelle *partition* d'un compact quarrable σ toute représentation de ce compact sous forme de réunion d'une famille finie de compacts quarrables dont les intérieurs sont disjoints deux à deux. Désignons les partitions par T, T', etc. Ainsi donc, la partition T d'un compact quarrable σ est une représentation de σ sous la forme $\sigma = \bigcup\limits_{k=1}^{n} \sigma_k$ où σ_k sont des compacts quarrables et $]\sigma_k[\cap]\sigma_l[= \varnothing$ pour $k \neq l$. Les compacts σ_k seront appelés *éléments* de T et l'aire de σ_k sera notée $\Delta\sigma_k$.

Désignons l'ensemble de toutes les partitions de σ par $\mathcal{T}_\sigma$. On peut munir $\mathcal{T}_\sigma$ de la relation d'ordre suivante : $T \leqslant T'$ si T' s'obtient à partir de T par une subdivision, i.e. chaque élément de T est soit un élément de T', soit la réunion de plusieurs éléments de T'. *Pour chaque couple de partitions* T' *et* T'' *de* $\mathcal{T}_\sigma$ *il existe une partition* T *de* $\mathcal{T}_\sigma$ *telle que* $T' \leqslant T$ *et* $T'' \leqslant T$.

En effet, soient σ_k' $(k = 1, \ldots, m)$ les éléments de T' et $\sigma_l''(l = 1, \ldots$ $\ldots, n)$ ceux de T''. Posons

$$A = \{(k, l) \mid \sigma_k' \cap \sigma_l'' \neq \varnothing\}, \qquad A_k = \{l \mid (k, l) \in A\} \ (k = 1, \ldots, m)$$

et

$$\sigma_{kl} = \sigma_k' \cap \sigma_l'' \quad \text{si } (k, l) \in A.$$

Ici σ_{kl} est un compact quarrable (σ_{kl} est un compact en tant qu'intersection de compacts, et une figure quarrable en tant qu'intersection de deux figures quarrables). Par hypothèse, $\sigma_{kl} \neq \varnothing$. Ceci étant, on a

$$1)\ \sigma = \left(\bigcup_{k=1}^{m} \sigma_k'\right) \cap \left(\bigcup_{l=1}^{n} \sigma_l''\right) = \bigcup_{k=1}^{m} \bigcup_{l=1}^{n} (\sigma_k' \cap \sigma_l'') = \bigcup_{(k,l) \in A} \sigma_{kl};$$

2) si (k_1, l_1) et (k_2, l_2) sont deux couples distincts de A, on a $]\sigma_{k_1 l_1}[\cap$ $\cap]\sigma_{k_2 l_2}[= \varnothing$. En effet, si par exemple $l_1 \neq l_2$, on obtient

$$]\sigma_{k_1 l_1}[\cap]\sigma_{k_2 l_2}[\subset]\sigma_{l_1}''[\cap]\sigma_{l_2}''[= \varnothing.$$

Ainsi donc, les compacts σ_{kl} forment une partition de σ ; notons-la T ;

3) $T' \leqslant T$. En effet, $\sigma_k' = \sigma_k' \cap \sigma = \sigma_k' \cap \bigcup_{l=1}^{n} \sigma_l'' = \bigcup_{l=1}^{n} (\sigma_k' \cap \sigma_l'') =$ $= \bigcup_{l \in A_k} \sigma_{kl}$. De façon analogue on démontre que $T'' \leqslant T$.

La partition T obtenue de façon mentionnée à partir de T' et T'' sera notée $T' \vee T''$.

Introduisons encore la notion de diamètre de la partition. On appelle *diamètre de l'ensemble E* dans un espace métrique la borne supérieure des distances entre les points de E si $E \neq \varnothing$, et 0 si $E = \varnothing$. Il est évident que l'ensemble est borné si et seulement si son diamètre est fini. En particulier, le diamètre de tout compact est fini. Par analogie aux partitions du segment, on appelle *diamètre de la partition* T d'un compact quarrable σ le diamètre maximal de ses éléments qu'on notera d_T ; *σ possède une partition de diamètre aussi petit que l'on veut*. Soit donné un ε. On subdivise le plan du compact σ en des carrés de diagonale inférieure à ε. Le compact σ n'a des points communs qu'avec une famille finie de ces carrés. En prenant les intersections de σ avec ces derniers pour éléments de partition, on obtient une partition T du compact σ avec $d_T < \varepsilon$.

2. Sommes inférieures et supérieures. Intégrales

Soit f une fonction arbitraire bornée sur un compact quarrable σ. Pour toute partition T de σ en $\sigma_1, \ldots, \sigma_n$, posons

$$m_k = \inf_{\sigma_k} f, \qquad M_k = \sup_{\sigma_k} f \qquad (k = 1, \ldots, n).$$

La fonction f étant bornée, m_k et M_k sont finis. Les sommes

$$\underline{s}_T = \sum_{k=1}^{n} m_k \Delta\sigma_k, \qquad \bar{s}_T = \sum_{k=1}^{n} M_k \Delta\sigma_k$$

seront respectivement appelées *sommes inférieure* et *supérieure* de f associées à T. Lorsque $f \geqslant 0$, elles ont un sens géométrique simple (fig. 33). Conformément au théorème 1 du § 2, $m_k \Delta\sigma_k$ est le volume d'un corps cylindrique $\mathcal{T}_k = \sigma_k \times [0, m_k]$ de base σ_k qui « supporte » le graphe de la fonction f. Si $k \neq l$, les corps $\mathcal{T}_k$ et $\mathcal{T}_l$ n'ont pas de points intérieurs communs puisque les projections des points intérieurs de ces corps sur le plan xOy sont des points intérieurs des bases et ces dernières n'ont pas de points intérieurs communs. Par conséquent, le corps « en escalier »

$$\mathcal{A}_T = \bigcup_{k=1}^{n} (\sigma_k \times [0, m_k])$$

« inscrit » dans le sous-graphe $\mathcal{F} = \{(x, y, z) \mid 0 \leqslant z \leqslant f(x, y)\}$ de la fonction f est cubable et $\underline{s}_T$ est son volume. De la même façon, $\bar{s}_T$ est le volume du corps « en escalier »

$$\mathcal{B}_T = \bigcup_{k=1}^{n} (\sigma_k \times [0, M_k])$$

« circonscrit » autour de $\mathcal{F}$.

Les sommes inférieures et supérieures possèdent les mêmes propriétés que dans le cas unidimensionnel.

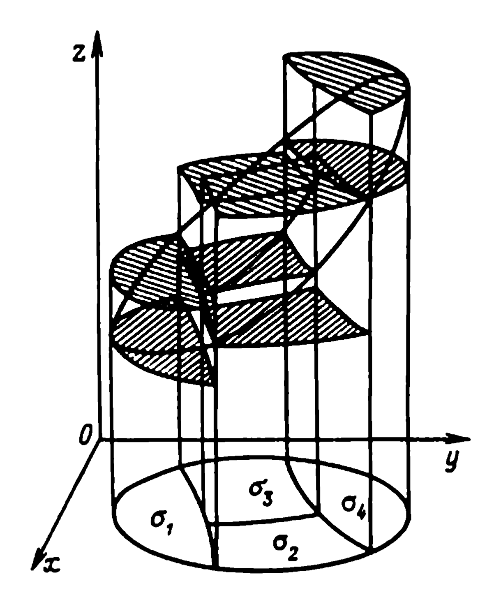


Fig. 33

1° $\underline{s}_T \leqslant \bar{s}_T$. Cela résulte directement des inégalités $m_k \leqslant M_k$ et $\Delta\sigma_k \geqslant 0$.

2° Si $T \leqslant T'$, on a $\underline{s}_T \leqslant \underline{s}_{T'}$ et $\bar{s}_T \geqslant \bar{s}_{T'}$. En effet, soient σ_k ($k = 1, \ldots$ $\ldots, n$) les éléments de la partition T. Puisque $T \leqslant T'$, ils peuvent être présentés sous la forme $\sigma_k = \bigcup_{l=1}^{l_k} \sigma_{kl}$ où σ_{kl} sont les éléments de la partition T'. Il peut arriver que les éléments différents σ_{k_1} et σ_{k_2} de T ont un terme commun $\sigma_{k_1 l_1} = \sigma_{k_2 l_2}$. Mais comme ils n'ont pas de points intérieurs communs, $]\sigma_{k_1 l_1}[\, = \,]\sigma_{k_2 l_2}[\, = \varnothing$ et c'est pourquoi $\Delta\sigma_{k_1 l_1} = \Delta\sigma_{k_2 l_2} = 0$. Par suite,

$$\underline{s}_{T'} = \sum_{k=1}^{n} \sum_{l=1}^{l_k} m_{kl} \Delta\sigma_{kl}, \qquad \bar{s}_{T'} = \sum_{k=1}^{n} \sum_{l=1}^{l_k} M_{kl} \Delta\sigma_{kl},$$

où $m_{kl} = \inf\limits_{\sigma_{kl}} f$, $M_{kl} = \sup\limits_{\sigma_{kl}} f$. Etant pris deux à deux, σ_{kl} n'ont pas de points communs, si bien que

$$\Delta\sigma_k = \sum_{l=1}^{l_k} \Delta\sigma_{kl} \qquad (k = 1, \ldots, n).$$

Vu que $\sigma_{kl} \subset \sigma_k$, on a $m_k \leqslant m_{kl}$ et $M_k \geqslant M_{kl}$. C'est pourquoi

$$\underline{s}_{T'} - \underline{s}_T = \sum_{k=1}^{n} \sum_{l=1}^{l_k} m_{kl} \Delta\sigma_{kl} - \sum_{k=1}^{n} m_k \Delta\sigma_k \geqslant \sum_{k=1}^{n} \sum_{l=1}^{l_k} m_k \Delta\sigma_{kl} -$$
$$- \sum_{k=1}^{n} m_k \Delta\sigma_k = \sum_{k=1}^{n} m_k \left(\sum_{l=1}^{l_k} \Delta\sigma_{kl} - \Delta\sigma_k \right) = 0.$$

On vérifie de façon analogue que $\bar{s}_T - \bar{s}_{T'} \geqslant 0$.

3° $\underline{s}_{T'} \leqslant \bar{s}_{T''}$ pour toutes partitions T' et T''. En effet, si $T' = T''$, c'est la propriété 1°. Lorsqu'il s'agit du cas général, posons $T = T' \vee T''$. Vu que $T' \leqslant T$ et $T'' \leqslant T$, les propriétés 2° et 1° impliquent

$$\underline{s}_{T'} \leqslant \underline{s}_T \leqslant \bar{s}_T \leqslant \bar{s}_{T''}.$$

Il résulte de la propriété 3° que :

1) la borne supérieure de la famille des sommes inférieures $(\underline{s}_T)_{T \in \mathcal{T}}$ est finie. On l'appelle *intégrale inférieure* de la fonction f sur σ et on la note $\underline{\iint_\sigma} f$;

2) la borne inférieure de la famille des sommes supérieures $(\bar{s}_T)_{T \in \mathcal{T}}$ est finie. Elle s'appelle *intégrale supérieure* de la fonction f sur σ et on la note $\overline{\iint_\sigma} f$;

3) $\underline{\iint_\sigma} f \leqslant \overline{\iint_\sigma} f$.

3. Intégrabilité. Intégrale double

DÉFINITION 1. La fonction f sur un compact σ quarrable est dite *intégrable* si elle est bornée sur σ et ses intégrales inférieure et supérieure sur σ coïncident. La valeur commune de ces intégrales est appelée *intégrale double de la fonction f sur σ* et elle est notée $\iint_\sigma f(x, y)\mathrm{d}x\mathrm{d}y$, ou $\iint_\sigma f \mathrm{d}_\sigma$, ou encore $\iint_\sigma f$.

EXEMPLE 1. L'intégrale double $\iint_\sigma c$, où c est une constante sur σ, existe et est égale à $c\Delta\sigma$. En effet, toutes les sommes inférieures et supérieures de c sur σ sont égales à $c\Delta\sigma$. En particulier,

$$\Delta\sigma = \iint_\sigma 1 = \iint_\sigma \mathrm{d}\sigma. \tag{1}$$

THÉORÈME 1. *La fonction $f: \mathbf{R}^2 \to \mathbf{R}$ est intégrable sur un compact quarrable σ si et seulement si elle est définie et bornée sur σ et*

$$(\forall \varepsilon > 0)(\exists T \in \mathscr{T}_\sigma)(\bar{s}_T - \underline{s}_T < \varepsilon).$$

COROLLAIRE 1 (interprétation géométrique de l'intégrale double). *Soit f une fonction réelle positive sur un compact quarrable σ. Si f est intégrable sur σ, son sous-graphe $\mathscr{F}^-$ est cubable et*

$$\text{vol}\, \mathscr{F}^- = \iint_\sigma f.$$

En effet, il a été noté au n° 2 que quelle que soit la partition $T \in \mathscr{T}_\sigma$, les solides en escalier $\mathscr{A}_T$ et $\mathscr{B}_T$ sont cubables et $\underline{s}_T = \text{vol}\, \mathscr{A}_T$, $\bar{s}_T = \text{vol}\, \mathscr{B}_T$. Conformément au théorème 1, il existe pour tout $\varepsilon > 0$ un $T \in \mathscr{T}_\sigma$ tel que $\bar{s}_T - \underline{s}_T < \varepsilon$. Donc, il existe pour tout $\varepsilon > 0$ des solides cubables $\mathscr{A}$ $(= \mathscr{A}_T)$ et $\mathscr{B}$ $(= \mathscr{B}_T)$ tels que $\mathscr{A} \subset \mathscr{F}^- \subset \mathscr{B}$ et $\text{vol}\, \mathscr{B} - \text{vol}\, \mathscr{A} < \varepsilon$. Cela signifie que $\mathscr{F}^-$ est cubable. Vu que

$$\underline{s}_T = \text{vol}\, \mathscr{A} \leqslant \text{vol}\, \mathscr{F}^- \leqslant \text{vol}\, \mathscr{B} = \bar{s}_T \quad \text{et} \quad \underline{s}_T \leqslant \iint_\sigma f = \underline{\iint_\sigma} f = \overline{\iint_\sigma} f \leqslant \bar{s}_T,$$

on a $0 \leqslant |\text{vol}\, \mathscr{F}^- - \iint_\sigma f| < \bar{s}_T - \underline{s}_T < \varepsilon$ et, puisque c'est vrai pour tout $\varepsilon > 0$, $\iint_\sigma f = \text{vol}\, \mathscr{F}^-$.

COROLLAIRE 2. *Le graphe Γ_f d'une fonction f définie et intégrable sur un compact quarrable σ est cubable et son volume est nul.*

Démonstration. Etant intégrable, f est bornée sur σ. Soit T une partition arbitraire de σ à éléments $\sigma_1, \ldots, \sigma_n$ et soient m_k, M_k, $\underline{s}_T$ et $\bar{s}_T$ des nombres introduits au n° 2. Posons

$$\mathscr{F}_T = \bigcup_{k=1}^{n} v_k \qquad \text{où} \qquad v_k = \sigma_k \times [m_k, M_k].$$

Il est évident que $\Gamma_f \subset \mathscr{F}_T$. En vertu du théorème 1 du § 2, les solides cylindriques v_k sont cubables et $\text{vol}\, v_k = (M_k - m_k)\Delta\sigma_k$. C'est pourquoi,

$\mathscr{T}_T$ est cubable et

$$\text{vol}\,\mathscr{T}_T \leqslant \sum_{k=1}^{n} \text{vol}\, v_k = \bar{s}_T - \underline{s}_T.$$

Le théorème 1 dit que $\bar{s}_T - \underline{s}_T$ peut devenir aussi petite que l'on veut grâce au choix de T. Donc, Γ_f est contenu dans les solides cubables de volume arbitrairement petit, i.e. il représente un solide cubable de volume nul.

THÉORÈME 2. *Toute fonction continue sur un compact quarrable est intégrable.*

Démonstration. Soit σ un compact quarrable. Si $\Delta\sigma = 0$, toute fonction bornée sur σ est intégrable : toutes les sommes $\underline{s}_T$ et $\bar{s}_T$ sont nulles, de sorte que $\iint_\sigma f$ existe et est égale à zéro. Soient $\Delta\sigma > 0$ et f une fonction continue sur σ. Comme σ est un compact, f est uniformément continue sur σ d'après le théorème 3.2.14, de sorte que

$$(\forall\varepsilon > 0)(\exists\delta > 0)(p,\ p' \in \sigma \text{ et } d(p,\ p') < \delta) \Rightarrow |f(p) - f(p')| < \frac{\varepsilon}{\Delta\sigma}.$$

Prenons une partition quelconque $T \in \mathscr{T}_\sigma$, avec $d_T < \delta$. Soient $\sigma_1, \ldots, \sigma_n$ ses éléments. Puisqu'ils sont des compacts et f est continue sur eux, il résulte du théorème 3.2.5 que

$$(\forall k \in [1,\ n])(\exists p_k,\ p_k' \in \sigma_k)\big(m_k = \inf_{\sigma_k} f = f(p_k),\ M_k = \sup_{\sigma_k} f = f(p_k')\big).$$

Ceci étant, $d(p_k,\ p_k') \leqslant \text{diam}\,\sigma_k \leqslant d_T < \delta$. Par suite,

$$\bar{s}_T - \underline{s}_T = \sum_{k=1}^{n} (M_k - m_k)\Delta\sigma_k = \sum_{k=1}^{n} [f(p_k') - f(p_k)]\Delta\sigma_k <$$

$$< \frac{\varepsilon}{\Delta\sigma} \sum_{k=1}^{n} \Delta\sigma_k = \varepsilon$$

et il ne reste qu'à appliquer le théorème 1.

REMARQUE. On a en effet démontré que pour une fonction continue f

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall T \in \mathscr{T}_\sigma)(d_T < \delta) \Rightarrow \bar{s}_T - \underline{s}_T < \varepsilon.$$

On verra plus loin que cela est juste pour toute fonction intégrable f.

§ 4. Propriétés fondamentales de l'intégrale double

1. Linéarité

(I) *Si f et g sont des fonctions bornées sur un compact quarrable σ, on a*

$$\underline{\iint}_\sigma f + \underline{\iint}_\sigma g \leqslant \underline{\iint}_\sigma (f + g) \leqslant \overline{\iint}_\sigma (f + g) \leqslant \overline{\iint}_\sigma f + \overline{\iint}_\sigma g.$$

Il en résulte directement que :

1° *Si les fonctions f et g sont intégrables sur un compact quarrable σ, leur somme est intégrable sur σ et*

$$\iint_\sigma (f+g) = \iint_\sigma f + \iint_\sigma g.$$

(II) *Si f est une fonction bornée sur un compact quarrable σ et $k \in \mathbf{R}$, on a*

$$\underline{\iint}_\sigma kf = \begin{cases} k \underline{\iint}_\sigma f \text{ pour } k \geqslant 0, \\ k \overline{\iint}_\sigma f \text{ pour } k < 0, \end{cases} \qquad \overline{\iint}_\sigma kf = \begin{cases} k \overline{\iint}_\sigma f \text{ pour } k \geqslant 0, \\ k \underline{\iint}_\sigma f \text{ pour } k < 0. \end{cases}$$

Il s'ensuit directement que :

2° *Si la fonction f est intégrable sur un compact quarrable σ, la fonction kf est intégrable sur σ pour tout k constant, et*

$$\iint_\sigma kf = k \iint_\sigma f.$$

On obtient de 1° et 2° par récurrence sur le nombre de termes :

3° *Si les fonctions $f_1, \ldots, f_n$ sont intégrables sur un compact quarrable σ, leur combinaison linéaire quelconque $k_1 f_1 + \ldots + k_n f_n$ est intégrable sur σ, et*

$$\iint_\sigma (k_1 f_1 + \ldots + k_n f_n) = k_1 \iint_\sigma f_1 + \ldots + k_n \iint_\sigma f_n.$$

En particulier,

$$\iint_\sigma (f_1 - f_2) = \iint_\sigma f_1 - \iint_\sigma f_2.$$

2. Additivité

(III) *Si la fonction f est définie et bornée sur des compacts quarrables σ' et σ'' n'ayant pas de points intérieurs communs, et $\sigma = \sigma' \cup \sigma''$, σ est un compact quarrable et*

$$\underline{\iint}_\sigma f = \underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f, \qquad \overline{\iint}_\sigma f = \overline{\iint}_{\sigma'} f + \overline{\iint}_{\sigma''} f.$$

Démonstration. En tant que réunion de deux figures quarrables, σ est une figure quarrable et en tant que réunion de deux compacts, σ est un compact. Soit T une partition de σ en $\sigma_1, \ldots, \sigma_n$. Posons

$$A' = \{k \mid \sigma_k \cap \sigma' \neq \varnothing\}, \qquad A'' = \{l \mid \sigma_l \cap \sigma'' \neq \varnothing\}$$

et $\sigma_k' = \sigma_k \cap \sigma'$ pour tout $k \in A'$, $\sigma_l'' = \sigma_l \cap \sigma''$ pour tout $l \in A''$.

Les ensembles σ_k' et σ_l'' sont quarrables (en tant qu'intersections de deux figures quarrables) et compacts (en tant qu'intersections des compacts). On

a de plus

$$\sigma' = \sigma' \cap \sigma = \sigma' \cap \bigcup_{k=1}^{n} \sigma_k = \bigcup_{k=1}^{n} (\sigma' \cap \sigma_k) = \bigcup_{k \in A'} \sigma'_k.$$

Enfin, $]\sigma'_{k_1}[\,\cap\,]\sigma'_{k_2}[\ \subset\ ]\sigma_{k_1}[\,\cap\,]\sigma_{k_2}[$ implique $]\sigma'_{k_1}[\,\cap\,]\sigma'_{k_2}[\ = \varnothing$ pour $k_1 \neq k_2$. Ainsi donc, σ'_k $(k \in A')$ forment une partition de σ' ; notons-la T'. De façon analogue, σ''_l $(l \in A'')$ forment une partition de σ'', on la note T''. Puisque $\sigma = \Big(\bigcup_{k \in A'} \sigma'_k\Big) \cup \Big(\bigcup_{l \in A'} \sigma''_l\Big)$ et $]\sigma'_k[\,\cap\,]\sigma''_l[\ (\subset\ ]\sigma_k[\,\cap\,]\sigma_l[) = \varnothing$, σ'_k et σ''_l forment ensemble une partition de σ ; notons-la $T' + T''$. Vu que $\sigma_k = \sigma_k \cap \sigma = \sigma_k \cap (\sigma' \cup \sigma'') = (\sigma_k \cap \sigma') \cup (\sigma_k \cap \sigma'')$, on a $T \leqslant T' + T''$. Or $\underline{s}_{T'+T''} = \underline{s}_{T'} + \underline{s}_{T''}$. Par conséquent, $\underline{s}_T \leqslant \underline{s}_{T'} + \underline{s}_{T''} \leqslant \underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f$, et en passant à la borne supérieure par rapport à T, on obtient

$$\underline{\iint}_{\sigma} f \leqslant \underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f. \tag{1}$$

D'autre part, vu que $\underline{\iint}_{\sigma'} f = \sup_{\mathscr{T}_{\sigma'}} \underline{s}_{T'}$ et $\underline{\iint}_{\sigma''} f = \sup_{\mathscr{T}_{\sigma''}} \underline{s}_{T''}$, il existe pour tout $\varepsilon > 0$ des partitions $T' \in \mathscr{T}_{\sigma'}$ et $T'' \in \mathscr{T}_{\sigma''}$ telles que

$$\underline{\iint}_{\sigma'} f < \underline{s}_{T'} + \frac{\varepsilon}{2}, \qquad \underline{\iint}_{\sigma''} f < \underline{s}_{T''} + \frac{\varepsilon}{2}.$$

Vu que σ' et σ'' n'ont pas de points intérieurs communs, T' et T'' forment ensemble une partition $T' + T'' \in \mathscr{T}_\sigma$. On a

$$\underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f < \underline{s}_{T'} + \underline{s}_{T''} + \varepsilon = \underline{s}_{T'+T''} + \varepsilon \leqslant \underline{\iint}_{\sigma} f + \varepsilon,$$

et comme ε est arbitraire, on obtient

$$\underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f \leqslant \underline{\iint}_{\sigma} f. \tag{2}$$

Il résulte de (1) et (2) que $\underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f = \underline{\iint}_{\sigma} f$. En remplaçant ici f par $-f$, on obtient en vertu de (II) que

$$\overline{\iint}_{\sigma'} f + \overline{\iint}_{\sigma''} f = -\underline{\iint}_{\sigma'} (-f) - \underline{\iint}_{\sigma''} (-f) = -\underline{\iint}_{\sigma} (-f) = \overline{\iint}_{\sigma} f.$$

La propriété (III) entraîne les assertions suivantes :

4° *Si la fonction f est intégrable sur des compacts quarrables σ' et σ'' n'ayant pas de points intérieurs communs, elle est intégrable sur la réunion σ de ces compacts et*

$$\iint_{\sigma} f = \iint_{\sigma'} f + \iint_{\sigma''} f.$$

5° *Si la fonction f est intégrable sur un compact quarrable σ, elle est intégrable sur tout compact quarrable σ' contenu dans σ.*

En effet, on peut admettre que $\sigma' \neq \sigma$; posons $\sigma'' = [\sigma \setminus \sigma']$. En tant que différence des figures quarrables, $\sigma \setminus \sigma'$ est quarrable (voir théorème 3 du § 1) et c'est pourquoi, son adhérence σ'' est un compact quarrable (voir théorème 5 du § 1). Vu que σ est un compact, il est fermé, ce qui implique $\sigma'' \subset [\sigma] = \sigma$. Il s'ensuit que $\sigma \supset \sigma' \cup \sigma'' \supset \sigma' \cup (\sigma \setminus \sigma') = \sigma$, de sorte que $\sigma = \sigma' \cup \sigma''$. Enfin, vu que $\sigma \setminus \,]\sigma'[$ est un ensemble fermé contenant $\sigma \setminus \sigma'$, on a $\sigma'' \subset \sigma \setminus \,]\sigma'[$ et par conséquent, $]\sigma''[$ a une intersection vide avec $]\sigma'[$. Donc, σ est la réunion des compacts quarrables σ' et σ'' qui n'ont pas de points intérieurs communs. Il vient alors d'après (III) :

$$\underline{\iint}_{\sigma'} f + \underline{\iint}_{\sigma''} f = \underline{\iint}_{\sigma} f = \overline{\iint}_{\sigma} f = \overline{\iint}_{\sigma'} f + \overline{\iint}_{\sigma''} f.$$

Puisque $\underline{\iint}_{\sigma'} f \leqslant \overline{\iint}_{\sigma'} f$ et $\underline{\iint}_{\sigma''} f \leqslant \overline{\iint}_{\sigma''} f$, il en résulte que $\underline{\iint}_{\sigma'} f = \overline{\iint}_{\sigma'} f$, i.e. f est intégrable sur σ'.

Il résulte par récurrence de 4° et 5° l'assertion suivante.

6° *Soient $\sigma_1, \ldots, \sigma_n$ des compacts quarrables dont les intérieurs sont disjoints deux à deux. La fonction f est intégrable sur leur réunion σ si et seulement si elle est intégrable sur chaque σ_k. Ceci étant, on a alors*

$$\iint_{\sigma} f = \sum_{k=1}^{n} \iint_{\sigma_k} f.$$

3. Intégration des inégalités. Théorème de la valeur moyenne

(IV) *Si les fonctions g et h sont définies et bornées sur un compact quarrable σ et $g \leqslant h$, on a*

$$\underline{\iint}_{\sigma} g \leqslant \underline{\iint}_{\sigma} h \qquad et \qquad \overline{\iint}_{\sigma} g \leqslant \overline{\iint}_{\sigma} h.$$

D é m o n s t r a t i o n. On entend par $\underline{s}_{\mathcal{T}}^{(f)}$ et $\overline{s}_{\mathcal{T}}^{(f)}$ les sommes inférieures et supérieures pour chaque fonction (bornée) f sur σ. Etant donné que $g \leqslant h$, on obtient $\underline{s}_{\mathcal{T}}^{(g)} \leqslant \underline{s}_{\mathcal{T}}^{(h)} \leqslant \underline{\iint}_{\sigma} h \ \left(= \sup\limits_{T \in \mathcal{T}_\sigma} \underline{s}_{\mathcal{T}}^{(h)}\right)$ pour toute partition $T \in \mathcal{T}_\sigma$ et c'est pourquoi on a aussi $\underline{\iint}_{\sigma} g \ \left(= \sup\limits_{T \in \mathcal{T}_\sigma} \underline{s}_{\mathcal{T}}^{(g)}\right) \leqslant \underline{\iint}_{\sigma} h$. L'inégalité pour les intégrales supérieures est obtenue de la dernière inégalité par substitution de $-g$ et $-h$ à g et h.

Il résulte directement de (IV) que :

7° *Si les fonctions g et h sont intégrables sur un compact quarrable σ et $g \leqslant h$, on a $\iint_{\sigma} g \leqslant \iint_{\sigma} h$.*

Vu que la constante nulle est intégrable et $\iint_\sigma 0 = 0$, on obtient en particulier :

8° *Si la fonction* $f \geqslant 0$ *est intégrable sur un compact quarrable* σ, *on a* $\iint_\sigma f \geqslant 0$.

8'° *Si la fonction* $f \geqslant 0$ *est intégrable sur un compact quarrable* σ *et est en outre continue et strictement positive en un point intérieur* p_0 *de* σ, *on a* $\iint_\sigma f > 0$.

En effet, $f \geqslant \frac{1}{2} f(p_0)$ sur un carré σ' de centre en p_0 contenu dans σ et distinct de σ. Soit $\sigma'' = [\sigma \setminus \sigma']$. Tout comme dans la démonstration de la propriété 5°, σ'' est un compact quarrable, $\sigma = \sigma' \cup \sigma''$ et $]\sigma'[\cap]\sigma''[= \varnothing$. D'après les propriétés 5°, 4°, 8°, 7° et la formule (1) du § 3, on a

$$\iint_\sigma f = \iint_{\sigma'} f + \iint_{\sigma''} f \geqslant \iint_{\sigma'} f \geqslant \iint_{\sigma'} \frac{1}{2} f(p_0) = \frac{1}{2} f(p_0) \Delta\sigma' > 0.$$

Il résulte directement de 8° et 3° que :

7'° *Si les fonctions g et h sont intégrables sur un compact quarrable* σ, $g \leqslant h$ *et* $g(p_0) < h(p_0)$ *en un point intérieur* p_0 *du compact* σ *où* $h - g$ *est continue, on a* $\iint_\sigma g < \iint_\sigma h$.

9° *Si la fonction f est intégrable sur un compact quarrable* σ *dont l'aire* $\Delta\sigma$ *est non nulle et si* $m \leqslant f(p) \leqslant M$ *pour tout* $p \in \sigma$, *on obtient*

$$m \leqslant \frac{1}{\Delta\sigma} \iint_\sigma f \leqslant M \tag{3}$$

(« *la moyenne intégrale d'une fonction a le même encadrement que la fonction* »). *Si de plus* $m < f(p_0) < M$ *en un point intérieur* p_0 *du compact* σ *où f est continue, la double inégalité de* (3) *est stricte.*

En effet, la formule (1) du § 3 et la propriété 7° impliquent $m\Delta\sigma = \iint_\sigma m \leqslant \iint_\sigma f \leqslant \iint_\sigma M = M\Delta\sigma$ et la division par $\Delta\sigma$ donne (3). La deuxième assertion découle de 7'° de façon analogue.

10° (THÉORÈME DE LA VALEUR MOYENNE). *La moyenne intégrale de la fonction continue f sur un compact quarrable connexe* σ *est égale à la valeur de f en un point de ce compact.*

D é m o n s t r a t i o n. Quand on parle de la moyenne intégrale sur σ, on suppose déjà que $\Delta\sigma \neq 0$. Vu que f est continue et σ est connexe, le théorème 3.1.1 implique que $f(\sigma)$ est un ensemble connexe dans **R**, i.e. est un intervalle. Etant donné que σ est un compact, cet intervalle est un compact d'après le théorème 3.2.6, i.e. est un segment d'après le théorème 3.2.13. Ainsi donc,

$f(\sigma) = [m, M]$. Il découle de la propriété 9° que $\frac{1}{\Delta\sigma} \iint_\sigma f \in [m, M]$ et par suite, $\frac{1}{\Delta\sigma} \iint_\sigma f$ est égale à la valeur de la fonction f en un point $p \in \sigma$.

4. Intégrabilité du module et du produit

11° *Si la fonction f est intégrable sur un compact quarrable σ, son module est intégrable sur σ et*

$$\left|\iint_\sigma f\right| \leqslant \iint_\sigma |f|.$$

12° *Si les fonctions f et g sont intégrables sur un compact quarrable σ, leur produit est également intégrable sur σ.*

5. Fonctions dont les points de discontinuité forment une figure d'aire nulle

13° *La fonction f bornée sur un compact quarrable σ dont les points de discontinuité forment une figure d'aire nulle, est intégrable sur σ.*

En effet, on peut évidemment poser $\Delta\sigma > 0$. Supposons que $|f| < C$ sur σ et soient Ψ l'ensemble des points de discontinuité de la fonction f et $\Phi = [\Psi]$. En vertu du théorème 4 du § 1, Φ est quarrable et aire $\Phi = 0$. C'est pourquoi, il existe pour tout $\varepsilon > 0$ un $n \in \mathbf{N}$ tel que

$$\text{aire } B_n(\Phi) < \min \{\varepsilon/(2C), \Delta\sigma\}.$$

Posons $\sigma' = \sigma \cap B_n(\Phi)$ et $\sigma'' = [\sigma \setminus \sigma']$. Comme aire $B_n(\Phi) < \Delta\sigma$, on a $\sigma' \neq \sigma$, de sorte que $\sigma'' \neq 0$. En tant qu'intersection des compacts quarrables, σ' est un compact quarrable et puisque $\sigma' \subset B_n(\Phi)$, on peut écrire $\Delta\sigma' < \frac{\varepsilon}{2C}$. Tout comme dans la démonstration de la propriété 5°, on voit que σ'' est un compact quarrable, $\sigma = \sigma' \cup \sigma''$ et $]\sigma'[\cap]\sigma''[= \varnothing$. Par conséquent, conformément à (III),

$$\overline{\iint_\sigma} f = \overline{\iint_{\sigma'}} f + \overline{\iint_{\sigma''}} f \quad \text{et} \quad \underline{\iint_\sigma} f = \underline{\iint_{\sigma'}} f + \underline{\iint_{\sigma''}} f, \tag{4}$$

d'où

$$\overline{\iint_\sigma} f - \underline{\iint_\sigma} f = \left(\overline{\iint_{\sigma'}} f - \underline{\iint_{\sigma'}} f\right) + \left(\overline{\iint_{\sigma''}} f - \underline{\iint_{\sigma''}} f\right). \tag{5}$$

Vu que $\sigma \setminus \sigma' = \sigma \setminus B_n(\Phi) \subset \sigma \setminus]B_n(\Phi)[$ et $\sigma \setminus]B_n(\Phi)[$ est fermé, $\sigma'' \subset \sigma \setminus]B_n(\Phi)[$. On démontre que $\Phi \subset]B_n(\Phi)[$. Par conséquent, $\sigma'' \subset \subset \sigma \setminus \Phi \subset \sigma \setminus \Psi$, de sorte que f est continue sur σ, et donc $\overline{\iint_{\sigma''}} f = \underline{\iint_{\sigma''}} f$

d'après le théorème 2 du § 3. C'est pourquoi, il résulte de (5) que

$$\overline{\iint_\sigma} f - \underline{\iint_\sigma} f \leqslant \left|\overline{\iint_{\sigma'}} f\right| + \left|\underline{\iint_{\sigma'}} f\right| .$$

Puisque $-C \leqslant f \leqslant C$, il vient d'après (IV)

$$-C\Delta\sigma' = \overline{\iint_{\sigma'}} (-C) \leqslant \overline{\iint_{\sigma'}} f \leqslant \overline{\iint_{\sigma'}} C = C\Delta\sigma',$$

de sorte que $\left|\overline{\iint_{\sigma'}} f\right| < C\dfrac{\varepsilon}{2C} = \dfrac{\varepsilon}{2}$ et de façon analogue $\left|\underline{\iint_{\sigma'}} f\right| <$ $< \dfrac{\varepsilon}{2}$. Par suite,

$$(0 \leqslant)\ \overline{\iint_\sigma} f - \underline{\iint_\sigma} f < \varepsilon,$$

et vu que ε est arbitraire, $\overline{\iint_\sigma} f = \underline{\iint_\sigma} f$, i.e. f est intégrable.

14° *La fonction f bornée sur un compact quarrable σ, qui n'est distincte de zéro que sur une figure Ψ d'aire nulle, est intégrable sur σ et son intégrale sur σ est nulle.*

En effet, soient σ' et σ'' des compacts quarrables introduits dans la démonstration de la propriété 13°. Vu que $\sigma' \subset \sigma \setminus \Psi$, on a $f = 0$ sur σ''. Il s'ensuit que $\overline{\iint_{\sigma''}} f = \underline{\iint_{\sigma''}} f = 0$ et donc, en vertu de (4), $\overline{\iint_\sigma} f = \overline{\iint_{\sigma'}} f$ et $\underline{\iint_\sigma} f = \underline{\iint_{\sigma'}} f$. Or on a vu dans la démonstration de la propriété 13° que les seconds membres de ces égalités peuvent devenir aussi petits que l'on veut grâce à un choix convenable de σ'. Par conséquent, $\overline{\iint_\sigma} f = \underline{\iint_\sigma} f = 0$, i.e. f est intégrable sur σ et $\iint_\sigma f = 0$.

14'° *Soient f et g des fonctions bornées sur un compact quarrable σ. Si f est intégrable sur σ et g n'est distincte de f que sur un ensemble Ψ d'aire nulle, g est intégrable sur σ et $\iint_\sigma g = \iint_\sigma f$.*

En effet, vu que la fonction $g - f$ est bornée sur σ et distincte de zéro seulement sur Ψ, elle est intégrable sur σ et $\iint_\sigma (g - f) = 0$ d'après la propriété 14°. Or $g = f + (g - f)$. Par conséquent, en vertu de la propriété 1°, g est intégrable sur σ et

$$\iint_\sigma g = \iint_\sigma f + \iint_\sigma (g - f) = \iint_\sigma f.$$

6. Intégrales inférieure et supérieure en tant que limites des sommes inférieures et supérieures. Deuxième critère de l'intégrabilité

(V) *Soit f une fonction bornée sur un compact quarrable σ. Alors*

$$\underline{\iint_\sigma} f = \lim_{d_T \to 0} \underline{s}_T, \qquad \overline{\iint_\sigma} f = \lim_{d_T \to 0} \bar{s}_T,$$

i.e.

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall T \in \mathscr{T}_\sigma)(d_T < \delta) \Rightarrow \left(\underline{\iint_\sigma} f - \underline{s}_T < \varepsilon \text{ et } \bar{s}_T - \overline{\iint_\sigma} f < \varepsilon\right).$$

Démonstration. Démontrons que

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall T \in \mathscr{T}_\sigma)(d_T < \delta) \Rightarrow \left(\bar{s}_T - \overline{\iint_\sigma} f\right) + \left(\underline{\iint_\sigma} f - \underline{s}_T\right) < \varepsilon,$$

i.e.

$$\bar{s}_T - \underline{s}_T < \overline{\iint_\sigma} f - \underline{\iint_\sigma} f + \varepsilon.$$

Cette relation est évidente pour $\Delta\sigma = 0$ ou $f \equiv 0$. Soient $\Delta\sigma > 0$ et $f \not\equiv 0$ sur σ, de sorte que $C = \sup_\sigma |f| > 0$. Vu que f est bornée parce qu'intégrable, on a $C < +\infty$. Prenons un $\varepsilon > 0$ arbitraire. Comme

$$\overline{\iint_\sigma} f = \inf_{T \in \mathscr{T}_\sigma} \bar{s}_T \qquad \text{et} \qquad \underline{\iint_\sigma} f = \sup_{T \in \mathscr{T}_\sigma} \underline{s}_T,$$

il existe des partitions T'_ε, $T''_\varepsilon \in \mathscr{T}_\sigma$ telles que

$$\bar{s}_{T'_\varepsilon} - \overline{\iint_\sigma} f < \frac{\varepsilon}{4} \qquad \text{et} \qquad \underline{\iint_\sigma} f - \underline{s}_{T''_\varepsilon} < \frac{\varepsilon}{4}.$$

Soit $T_\varepsilon = T'_\varepsilon \vee T''_\varepsilon$ (voir n° 1 du § 3). En prenant en considération la propriété 2° des sommes inférieures et supérieures, établie au n° 2 du § 3, on obtient

$$\bar{s}_{T_\varepsilon} - \overline{\iint_\sigma} f < \frac{\varepsilon}{4} \qquad \text{et} \qquad \underline{\iint_\sigma} f - \underline{s}_{T_\varepsilon} < \frac{\varepsilon}{4},$$

d'où

$$\bar{s}_{T_\varepsilon} - \underline{s}_{T_\varepsilon} < \overline{\iint_\sigma} f - \underline{\iint_\sigma} f + \frac{\varepsilon}{2}. \tag{6}$$

Soient $\sigma_1^\varepsilon, \ldots, \sigma_m^\varepsilon$ les éléments de la partition T_ε et soient $m_k^\varepsilon = \inf_{\sigma_k^\varepsilon} f$, $M_k^\varepsilon = \sup_{\sigma_k^\varepsilon} f$ et $\Gamma = \bigcup_{k=1}^m \operatorname{Fr} \sigma_k^\varepsilon$. Vu que σ_k^ε sont quarrables, on obtient d'après le théorème 2 du § 1 et le lemme du § 1 que Γ est quarrable et aire

$\Gamma = 0$. Pour cette raison, il existe un $n \in \mathbf{N}$ tel que $B_n(\Gamma) < \frac{\varepsilon}{4}$. On démontre que $\Gamma \subset]B_n(\Gamma)[$, de sorte que $\Gamma \cap \operatorname{Fr} B_n(\Gamma) = \varnothing$. Or Γ et $B_n(\Gamma)$ sont des compacts d'après les théorèmes 3.3.3 et 3.2.4. C'est pourquoi, la distance δ entre Γ et $\operatorname{Fr} B_n(\Gamma)$ est strictement positive d'après la propriété 3.2.5, 5° et, par suite, $U(\Gamma; \delta) \subset B_n(\Gamma)$ en vertu du théorème 3.3.4. Montrons que

$$d_T < \delta \Rightarrow \bar{s}_T - \underline{s}_T < \overline{\iint_\sigma} f - \underline{\iint_\sigma} f + \varepsilon.$$

Etant donné $d_T < \delta$, soient $\sigma_1, \ldots, \sigma_n$ les éléments de la partition T, $m_i = \inf\limits_{\sigma_i} f$, $M_i = \sup\limits_{\sigma_i} f$. Puisque $\sigma_i \subset \sigma = \bigcup\limits_{k=1}^{m} \sigma_k^\varepsilon$, il existe pour tout entier $i \in [1, n]$ un $k \in [1, m]$ tel que $\sigma_i \cap \sigma_k^\varepsilon \neq \varnothing$, i.e. il existe un point $p' \in \sigma_i \cap \sigma_k^\varepsilon$. Alors $\sigma_i \subset \sigma_k^\varepsilon$ ou $\sigma_i \subset B_n(\Gamma)$. En effet, si $\sigma_i \not\subset \sigma_k^\varepsilon$, i.e. s'il existe un point $p'' \in \sigma_i \setminus \sigma_k^\varepsilon$, on trouve d'après le théorème 3.3.2 un point $p_0 \in [p', p''[\cap \operatorname{Fr} \sigma_k^\varepsilon$ puisque $p' \in \sigma_k^\varepsilon$ et $p'' \notin \sigma_k^\varepsilon$. On a $p_0 \in [p', p'']$, i.e. $p_0 = (1 - \lambda)p' + \lambda p''$ où $\lambda \in [0, 1]$, si bien que pour tout point p du plan on obtient

$$\|p - p_0\| = \|(1 - \lambda)(p - p') + \lambda(p - p'')\| \leqslant$$
$$\leqslant (1 - \lambda)\|p - p'\| + \lambda\|p - p''\| \leqslant \max\{\|p - p'\|, \|p - p''\|\}$$

(fig. 34). Puisque $p', p'' \in \sigma_i$, il en résulte que quel que soit le point $p \in \sigma_i$,

$$\|p - p_0\| \leqslant \max\{\|p - p'\|, \|p - p''\|\} \leqslant \operatorname{diam} \sigma_i \leqslant d_T < \delta$$

et, par conséquent,

$$\sigma_i \subset U(p_0; \delta) \subset U(\operatorname{Fr} \sigma_k^\varepsilon; \varepsilon) \subset U(\Gamma; \delta) \subset B_n(\Gamma).$$

Maintenant,

$$\bar{s}_T - \underline{s}_T = \Sigma_1 + \Sigma_2, \text{ où } \Sigma_1 = \sum_{\sigma_i \not\subset B_n(\Gamma)} (M_i - m_i)\Delta\sigma_i,$$

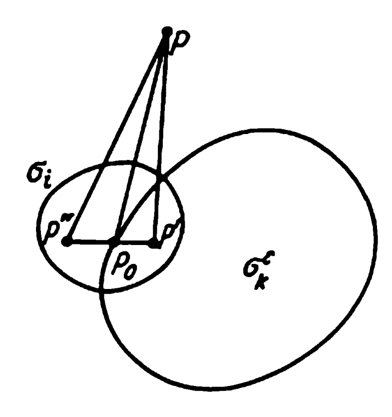


Fig. 34

(7)
$$\Sigma_2 = \sum_{\sigma_j \subset B_n(\Gamma)} (M_j - m_j)\Delta\sigma_j.$$

Si $\sigma_i \not\subset B_n(\Gamma)$, on a $\sigma_i \subset \sigma_k^\varepsilon$ pour un k d'après ce qui a été démontré. C'est pourquoi

$$\Sigma_1 \leqslant \sum_{k=1}^{m} \sum_{\sigma_i \subset \sigma_k^\varepsilon} (M_i - m_i)\Delta\sigma_i.$$

Or $M_i - m_i \leqslant M_k^\varepsilon - m_k^\varepsilon$ si $\sigma_i \subset \sigma_k^\varepsilon$. Par suite,

(8)
$$\Sigma_1 \leqslant \sum_{k=1}^{m} (M_k^\varepsilon - m_k^\varepsilon) \sum_{\sigma_i \subset \sigma_k^\varepsilon} \Delta\sigma_i \leqslant \sum_{k=1}^{m} (M_k^\varepsilon - m_k^\varepsilon)\Delta\sigma_k^\varepsilon = \bar{s}_{T_\varepsilon} - \underline{s}_{T_\varepsilon}.$$

D'autre part, comme $|M_j| \leqslant C$ et $|m_j| \leqslant C$, on a

(9)
$$\Sigma_2 = 2C \sum_{\sigma_j \subset B_n(\Gamma)} \Delta\sigma_j \leqslant 2C \text{ aire } B_n(\Gamma) < \frac{\varepsilon}{2}.$$

Il résulte alors de (7), (8), (9) et (6) que

$$\bar{s}_T - \underline{s}_T < \bar{s}_{T_\varepsilon} - \underline{s}_{T_\varepsilon} + \frac{\varepsilon}{2} < \overline{\iint_\sigma} f - \underline{\iint_\sigma} f + \varepsilon.$$

La propriété (V) entraîne immédiatement :

15° *La fonction f bornée sur un compact quarrable σ est intégrable si et seulement si*

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall T \in \mathscr{T}_\sigma)(d_T < \delta) \Rightarrow \bar{s}_T - \underline{s}_T < \varepsilon.$$

7. Intégrale double de Riemann

Soit f une fonction sur un compact quarrable σ. Considérons une partition T de σ en $\sigma_1, \ldots, \sigma_n$ et choisissons dans chaque σ_k un point ξ_k ; désignons l'ensemble de ces points par Ξ et notons (T, Ξ) la *partition T de σ munie des points choisis*. On appelle *somme de Riemann* de f attachée à (T, Ξ) la somme

$$s_{T,\Xi} = \sum_{k=1}^{m} f(\xi_k)\Delta\sigma_k.$$

DÉFINITION 1. Soit f une fonction définie sur un compact quarrable σ. On dit qu'elle est *intégrable au sens de Riemann* sur σ si elle est bornée sur σ et si $s_{T,\Xi}$ tend vers un nombre J lorsque $d_T \to 0$, i.e.

(10)
$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall (T, \Xi))(d_T < \delta) \Rightarrow |J - s_{T,\Xi}| < \varepsilon.$$

Le nombre J est alors appelé *intégrale double de Riemann* de la fonction f sur σ.

REMARQUE. Dans le cas unidimensionnel, la condition de f bornée est superflue dans la définition de l'intégrabilité au sens de Riemann puisque cette condition découle de (10). Dans le cas bidimensionnel, la situation est autre : il existe des compacts quarrables infinis d'aire nulle sur lesquels toute fonction, quoique non bornée, vérifie la condition (10) avec $J = 0$.

Tout comme dans le cas unidimensionnel, on démontre la propriété suivante :

16° *La fonction f sur un compact quarrable σ est intégrable au sens de Riemann si et seulement si elle est bornée et ses intégrales inférieure et supérieure sur σ coïncident. Leur valeur commune est alors égale à l'intégrale de Riemann de la fonction f sur σ.*

§ 5. Calcul de l'intégrale double par intégration répétée

Le calcul de l'intégrale double pour certains domaines d'intégration peut être ramené au calcul de deux intégrales définies.

1. Calcul de l'intégrale double sur un rectangle dont les côtés sont parallèles aux axes de coordonnées

THÉORÈME 1. *Soit σ un rectangle fermé borné à gauche et à droite par les droites $x = a$ et $x = b$ et par les droites $y = c$ et $y = d$ en bas et en haut. Si une fonction $f(x, y)$ est intégrable sur σ et si pour tout $x \in [a, b]$ fixé elle est intégrable sur le segment $[c, d]$ en tant que fonction de y, la fonction $g(x) = \int_c^d f(x, y)\mathrm{d}y$ est intégrable sur $[a, b]$ et*

$$\iint_\sigma f = \int_a^b \left(\int_c^d f(x, y)\mathrm{d}y\right)\mathrm{d}x.$$

Au lieu de $\int_a^b \left(\int_c^d f(x, y)\mathrm{d}y\right)\mathrm{d}x$, on écrira aussi $\int_a^b\int_c^d f(x, y)\mathrm{d}y\mathrm{d}x$.

La fig. 35 donne une interprétation géométrique de ce théorème (pour une fonction $f \geqslant 0$) : $g(x)$ est l'aire de la section du sous-graphe de f par le plan $x =$ const (perpendiculaire au plan $z = 0$), de sorte que le théorème affirme que le volume du sous-graphe est l'intégrale de l'aire de la section (tout comme l'aire du sous-graphe de la fonction positive d'une variable est l'intégrale de la longueur de l'ordonnée).

Démonstration du théorème 1. Soient T_1 et T_2 les subdivisions de $[a, b]$ et $[c, d]$ par des points x_i ($i \in [0, m]$) et y_j ($j \in [0, n]$) : $a = x_0 < x_1 < \ldots < x_{m-1} < x_m = b$, $c = y_0 < y_1 < \ldots < y_{n-1} < y_n = d$. Désignons par T_{12} la partition de σ en rectangles

$$\sigma_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}] \qquad (i \in [0, m-1], \qquad j \in [0, n-1])$$

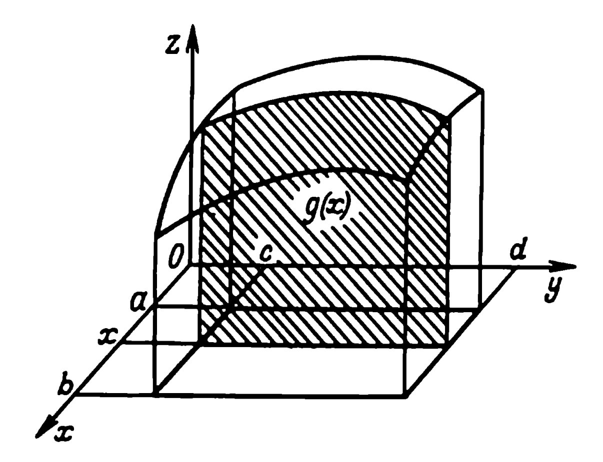


Fig. 35

et soient $m_{ij} = \inf_{\sigma_{ij}} f$, $M_{ij} = \sup_{\sigma_{ij}} f$, $m_i = \inf_{[x_i, x_{i+1}]} g$, $M_i = \sup_{[x_i, x_{i+1}]} g$. Comme $m_{ij} \leqslant f(x, y) \leqslant M_{ij}$ pour tout $(x, y) \in \sigma_{ij}$, on a

$$m_{ij}\Delta y_j \leqslant \int_{y_j}^{y_{j+1}} f(x, y)\mathrm{d}y \leqslant M_{ij}\Delta y_j$$

pour tout $x \in [x_i, x_{i+1}]$. C'est pourquoi,

$$\sum_{j=0}^{n-1} m_{ij}\Delta y_j \leqslant \sum_{j=0}^{n-1} \int_{y_j}^{y_{j+1}} f(x, y)\mathrm{d}y = g(x) \leqslant \sum_{j=0}^{n-1} M_{ij}\Delta y_j$$

pour tout $x \in [x_i, x_{i+1}]$, d'où

$$\sum_{j=0}^{n-1} m_{ij}\Delta y_j \leqslant m_i \leqslant M_i \leqslant \sum_{j=0}^{n-1} M_{ij}\Delta y_j.$$

En multipliant chaque membre de ces inégalités par Δx_i et en faisant la sommation en i de 0 à $m - 1$, on obtient

$$(1)\quad \underline{s}_{T_{12}}^{(f)} = \sum_{i=0}^{m-1} \left(\sum_{j=0}^{n-1} m_{ij}\Delta y_j\right)\Delta x_i \leqslant \underline{s}_{T_1}^{(g)} \leqslant \overline{s}_{T_1}^{(g)} \leqslant$$
$$\leqslant \sum_{i=0}^{m-1} \left(\sum_{j=0}^{n-1} M_{ij}\Delta y_j\right)\Delta x_i = \overline{s}_{T_{12}}^{(f)}.$$

Vu que f est intégrable sur σ, on a

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall T \in \mathscr{T}_\sigma)(d_T < \delta) \Rightarrow \overline{s}_T^{(f)} - \underline{s}_T^{(f)} < \varepsilon$$

(voir propriété 15° du § 4). Or diam $T_{12} = \max_{i,j} \sqrt{\Delta x_i^2 + \Delta y_j^2} < \delta$ si $d_{T_1} < \delta/\sqrt{2}$ et $d_{T_2} < \delta/\sqrt{2}$. Par conséquent,

$$\overline{s}_{T_{12}}^{(f)} - \underline{s}_{T_{12}}^{(f)} < \varepsilon$$

et à plus forte raison en vertu de (1)

$$\overline{S}_{T_1}^{(g)} - \underline{S}_{T_1}^{(g)} < \varepsilon.$$

ε étant arbitraire, il en résulte que g est intégrable sur $[a, b]$ et, en vertu des inégalités (1),

$$\underline{S}_{T_{12}}^{(f)} \leqslant \int_a^b g \leqslant \overline{S}_{T_{12}}^{(f)}.$$

Comme

$$\underline{S}_{T_{12}}^{(f)} \leqslant \iint_\sigma f \leqslant \overline{S}_{T_{12}}^{(f)},$$

on a $\left| \iint_\sigma f - \int_a^b g \right| < \varepsilon$ pour tout $\varepsilon > 0$ et, par suite, $\iint_\sigma f = \int_a^b g$.

Le théorème suivant est démontré de façon analogue.

THÉORÈME 1'. *Soit σ un rectangle fermé borné à gauche et à droite par les droites $x = a$ et $x = b$ et par les droites $y = c$ et $y = d$ en bas et en haut. Si une fonction $f(x, y)$ est intégrable sur σ et si pour tout $y \in [c, d]$ donné elle est intégrable sur le segment $[a, b]$ en tant que fonction de x, la fonction*

$$h(y) = \int_a^b f(x, y)\mathrm{d}x$$

est intégrable sur $[c, d]$ et

$$\iint_\sigma f = \int_c^d \left(\int_a^b f(x, y)\mathrm{d}x\right)\mathrm{d}y.$$

L'expression $\int_a^b \left(\int_c^d f(x, y)\mathrm{d}x\right)\mathrm{d}y$ sera aussi notée $\int_c^d\int_a^b f(x, y)\mathrm{d}x\mathrm{d}y$.

Remarquons que si f est continue sur σ, les hypothèses des deux théorèmes sont vérifiées et, donc, on est libre de choisir l'argument par lequel on commence l'intégration.

EXEMPLE 1. Il faut intégrer la fonction $f(x, y) = \dfrac{x}{(1 + x^2 + y^2)^{3/2}}$ sur le carré $\sigma = [0, 1] \times [0, 1]$. Il est plus rationnel de l'intégrer d'abord par rapport à x (en prenant y pour une constante) et ensuite par rapport à y. Procédant à la substitution $1 + x^2 + y^2 = u$, on obtient

$$\int_0^1 \frac{x}{(1 + x^2 + y^2)^{3/2}}\mathrm{d}x = \frac{1}{2}\int_{1+y^2}^{2+y^2} \frac{\mathrm{d}u}{u^{3/2}} = \frac{u^{-1/2}}{-1/2}\Bigg|_{1+y^2}^{2+y^2} = \frac{1}{\sqrt{1 + y^2}} - \frac{1}{\sqrt{2 + y^2}},$$

d'où

$$\int_0^1\int_0^1 \frac{x}{(1+x^2+y^2)^{3/2}}\,\mathrm{d}x\mathrm{d}y = \int_0^1 \frac{\mathrm{d}y}{\sqrt{1+y^2}} - \int_0^1 \frac{\mathrm{d}y}{\sqrt{2+y^2}}.$$

Il découle de l'intégrale indéfinie $\int \frac{\mathrm{d}u}{\sqrt{u^2+1}} = \ln(u + \sqrt{u^2+1}) + C$ que $\int_0^1 \frac{\mathrm{d}y}{\sqrt{1+y^2}} =$ $= \ln(1+\sqrt{2})$. La substitution $y = u\sqrt{2}$ dans l'intégrale $\int_0^1 \frac{\mathrm{d}y}{\sqrt{2+y^2}}$ donne $\int_0^{1/\sqrt{2}} \frac{\mathrm{d}u}{\sqrt{u^2+1}} =$ $= \ln \frac{1+\sqrt{3}}{\sqrt{2}}$. Ainsi donc, $\int_0^1\int_0^1 \frac{x}{(1+x^2+y^2)^{3/2}}\,\mathrm{d}x\mathrm{d}y = \ln \frac{(1+\sqrt{2})\sqrt{2}}{1+\sqrt{3}}$.

EXERCICE 1. Intégrer cette fonction d'abord par rapport à y et puis par rapport à x.

2. Calcul de l'intégrale double sur un domaine compris entre deux graphes

THÉORÈME 2. *Soient* φ_1 *et* φ_2 *des fonctions continues sur le segment* $[a, b]$ *et* $\varphi_1(x) \leqslant \varphi_2(x)$ *pour tout* $x \in [a, b]$. *Alors*

$$\sigma = \{(x, y) \in \mathbf{R}^2 \mid a \leqslant x \leqslant b \quad et \quad \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$$

est un compact quarrable. Si une fonction $f(x, y)$ *est intégrable sur* σ *et si pour tout* $x \in [a, b]$ *fixé elle est intégrable sur le segment* $[\varphi_1(x), \varphi_2(x)]$ *en tant que fonction de* y, *la fonction*

$$g(x) = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\mathrm{d}y$$

est intégrable sur $[a, b]$ *et*

$$\iint_\sigma f = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\mathrm{d}y\right)\mathrm{d}x$$

(fig. 36).

Démonstration. Il est aisé d'établir que σ est quarrable et borné. En outre, σ est fermé. En effet, si $(x_n, y_n) \in \sigma$ et $(x_n, y_n) \to (x, y)$, i.e. $x_n \to x$ et $y_n \to y$, on a

$$a \leqslant x_n \leqslant b \Rightarrow a \leqslant x \leqslant b$$

et, en vertu de la continuité des fonctions φ_1 et φ_2,

$$\varphi_1(x_n) \leqslant y_n \leqslant \varphi_2(x_n) \Rightarrow \varphi_1(x) \leqslant y \leqslant \varphi_2(x),$$

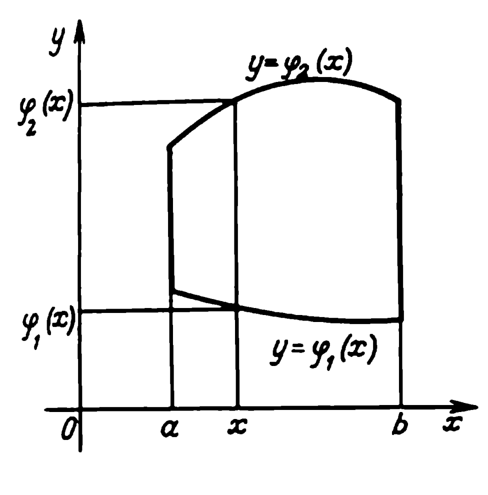


Fig. 36

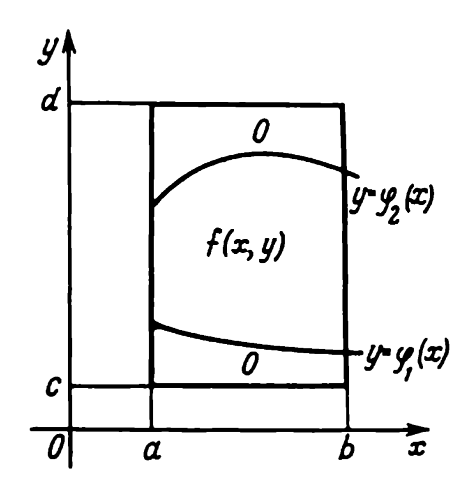


Fig. 37

de sorte que $(x, y) \in \sigma$. Ainsi donc, σ est un compact quarrable. En vertu de leur continuité, φ_1 et φ_2 sont bornées sur $[a, b]$. Soient c et d des nombres quelconques satisfaisant aux conditions $c < \varphi_1(x)$ et $d > \varphi_2(x)$ pour tout $x \in [a, b]$. Posons

$$\sigma^* = [a, b] \times [c, d],$$
$$f^*(x, y) = \begin{cases} f(x, y) & \text{sur } \sigma, \\ 0 \text{ sur } & \sigma^* \setminus \sigma \end{cases}$$

(fig. 37) ; f^* et σ^* vérifient les hypothèses du théorème 1, i.e. f^* est intégrable sur σ^* et $f^*(x, y)$ est intégrable sur $[c, d]$ pour tout $x \in [a, b]$ fixé. En effet, soit $\sigma' = [\sigma^* \setminus \sigma]$; σ' est un compact quarrable composé de deux parties

$$\sigma_1' = \{(x, y) \in \mathbf{R}^2 \mid a \leqslant x \leqslant b \quad \text{et} \quad c \leqslant y \leqslant \varphi_1(x)\},$$
$$\sigma_2' = \{(x, y) \in \mathbf{R}^2 \mid a \leqslant x \leqslant b \quad \text{et} \quad \varphi_2(x) \leqslant y \leqslant d\}.$$

La fonction f^* est intégrable sur σ puisqu'elle y coïncide avec f et elle est intégrable sur σ' en vertu de la propriété 14° de l'intégrale double (voir § 4) puisqu'elle peut y être distincte de zéro seulement sur les graphes des fonctions φ_1 et φ_2 et la réunion de ces graphes en tant qu'une partie de Frσ est une figure d'aire nulle (voir théorème 2 du § 1). Conformément aux propriétés 4° et 14° de l'intégrale double,

$$\iint_{\sigma^*} f^* = \iint_{\sigma} f^* + \iint_{\sigma'} f^* = \iint_{\sigma} f.$$

Enfin pour tout $x \in [a, b]$, on a

$$[c, d] = [c, \varphi_1(x)] \cup [\varphi_1(x), \varphi_2(x)] \cup [\varphi_2(x), d].$$

Ceci étant, $f^*(x, y)$ peut être distincte de zéro sur $[c, \varphi_1(x)]$ seulement au point $y = \varphi_1(x)$ et sur $[\varphi_2(x), d]$, seulement au point $y = \varphi_2(x)$, et elle est intégrable comme fonction de y sur le segment $[\varphi_1(x), \varphi_2(x)]$ car elle y coïncide avec $f(x, y)$. Par conséquent, quel que soit $x \in [a, b]$, $f^*(x, y)$ est intégrable sur $[c, d]$ en tant que fonction de y et

$$\int_c^d f^*(x, y)\mathrm{d}y = \int_c^{\varphi_1(x)} f^*(x, y)\mathrm{d}y + \int_{\varphi_1(x)}^{\varphi_2(x)} f^*(x, y)\mathrm{d}y + \int_{\varphi_2(x)}^{d} f^*(x, y)\mathrm{d}y =$$

$$= \int_{\varphi_1(x)}^{\varphi_2(x)} f^*(x, y)\mathrm{d}y.$$

En appliquant le théorème 1 à f^* et à σ^*, on obtient

$$\iint_\sigma f = \iint_{\sigma^*} f^* = \int_a^b\int_c^d f^*(x, y)\mathrm{d}y\mathrm{d}x = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\mathrm{d}y\mathrm{d}x.$$

On démontre de façon analogue le théorème suivant.

THÉORÈME 2′. *Soient ψ_1 et ψ_2 des fonctions continues sur le segment $[c, d]$ et $\psi_1(y) \leqslant \psi_2(y)$ pour tout $y \in [c, d]$. Alors*

$$\sigma = \{(x, y) \in \mathbf{R}^2 \mid c \leqslant y \leqslant d \quad \text{et} \quad \psi_1(y) \leqslant x \leqslant \psi_2(y)\}$$

est un compact quarrable. Si une fonction $f(x, y)$ est intégrable sur σ et si pout tout $y \in [c, d]$ fixé elle est intégrable sur le segment $[\psi_1(y), \psi_2(y)]$ comme fonction de x, la fonction

$$h(y) = \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\mathrm{d}x$$

est intégrable sur $[c, d]$ et

$$\iint_\sigma f = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\mathrm{d}x\mathrm{d}y$$

(fig. 38).

EXEMPLE 2. Soit $a > 0$, $b > 0$, $c > 0$,

$$\sigma = \left\{(x, y) \in \mathbf{R}^2 \mid -a \leqslant x \leqslant a, \ -b\sqrt{1 - \frac{x^2}{a^2}} \leqslant y \leqslant b\sqrt{1 - \frac{x^2}{a^2}}\right\}$$

et

$$f(x, y) = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}.$$

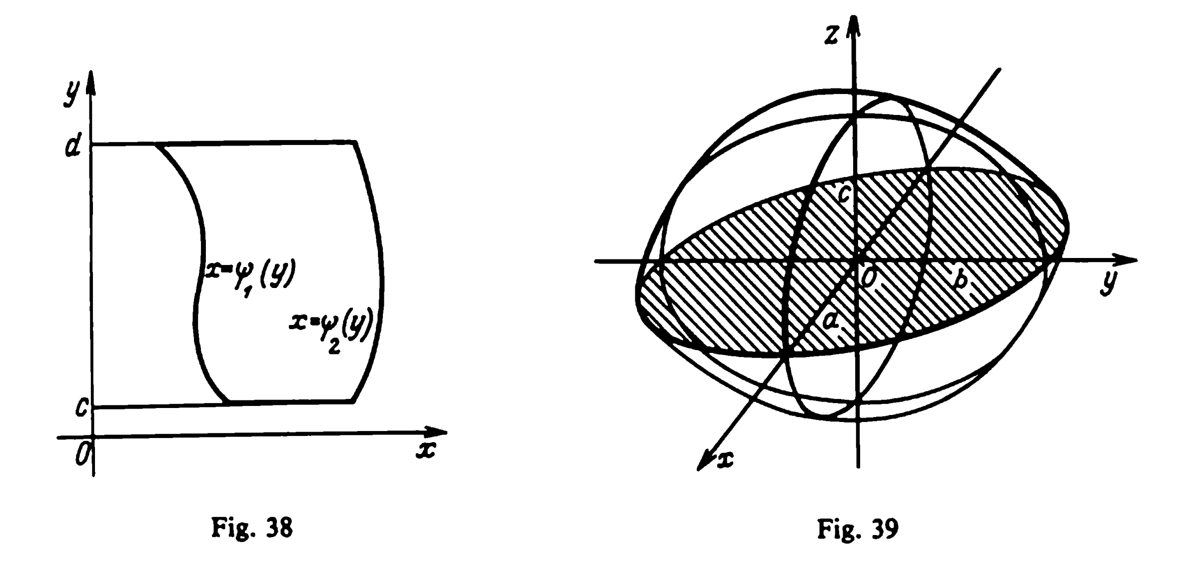


Fig. 38

Fig. 39

Il est aisé de voir que σ est l'ellipse

$$\left\{(x, y) \in \mathbf{R}^2 \,\middle|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}.$$

Pour cette raison, $D_f = \sigma$ et, d'après le théorème 2 (dont les hypothèses sont ici vérifiées),

$$\iint_\sigma f = \int_{-a}^{a} \int_{-b\sqrt{1-\frac{x^2}{a^2}}}^{b\sqrt{1-\frac{x^2}{a^2}}} c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\, \mathrm{d}y\mathrm{d}x.$$

Pour calculer l'intégrale intérieure (avec x fixé) on procède à une substitution :

$$y = b\sqrt{1 - \frac{x^2}{a^2}} \sin t.$$

Quand t parcourt le segment $[-\pi/2, \pi/2]$, y parcourt le segment $\left[-b\sqrt{1 - \frac{x^2}{a^2}}, b\sqrt{1 - \frac{x^2}{a^2}}\right]$. Ceci étant, on a

$$\mathrm{d}y = b\sqrt{1 - \frac{x^2}{a^2}} \cos t\mathrm{d}t, \quad \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} = \sqrt{1 - \frac{x^2}{a^2}} \cos t,$$

donc

$$\iint_\sigma f = \int_{-a}^{a} \int_{-\pi/2}^{\pi/2} bc\left(1 - \frac{x^2}{a^2}\right) \cos^2 t\mathrm{d}t\mathrm{d}x = $$

$$= bc \int_{-a}^{a} \left(\int_{-\pi/2}^{\pi/2} \cos^2 t\mathrm{d}t\right)\left(1 - \frac{x^2}{a^2}\right)\mathrm{d}x.$$

Vu que $\int_{-\pi/2}^{\pi/2} \cos^2 t\mathrm{d}t = \frac{\pi}{2}$, on obtient

$$\iint_\sigma f = \frac{\pi}{2} bc \int_{-a}^{a} \left(1 - \frac{x^2}{a^2}\right)\mathrm{d}x = \frac{2}{3}\pi abc.$$

Il est facile de voir que la moitié supérieure de l'ellipsoïde

$$\left\{(x, y, z) \middle| \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1\right\}$$

sert de sous-graphe de la fonction f (fig. 39). Puisque l'ellipsoïde est symétrique par rapport au plan xOy, il résulte du corollaire au théorème 1 du § 3 qu'il est cubable et que son volume V s'exprime en fonction des demi-axes a, b, c par la formule $V = \frac{4}{3}\pi abc$. Si $a = b = c$, il en résulte la cubabilité de la boule et la formule classique $V = \frac{4}{3}\pi a^3$ exprimant son volume par le rayon a.

CHAPITRE 12

CHANGEMENT DE VARIABLES DANS L'INTÉGRALE DOUBLE

§ 1. Applications régulières

Ce chapitre est consacré au problème de la transformation de l'intégrale double

$$\iint_{\sigma} F(x, y)\mathrm{d}x\mathrm{d}y$$

par changement de variables de la forme

$$x = \varphi(u, v), \ y = \psi(u, v).$$

Il s'avère rationnel d'envisager φ et ψ en tant que composantes de l'application d'un sous-ensemble ouvert du plan des coordonnées u, v dans le plan des coordonnées x, y.

1. Notion d'application régulière

DÉFINITION 1. Une application f: $\mathbf{R}^n \to \mathbf{R}^n$ définie sur un ensemble ouvert non vide D est dite *régulière* si : 1) f est continûment différentiable sur D, 2) $f'(w)$ est inversible pour tout $w \in D$ et 3) f est injective.

Si $f_1, \ldots, f_n$ sont les composantes de l'application f, la condition 1) de la définition 1 est équivalente, en vertu du théorème 9.3.5, à ce que les composantes ont partout des dérivées partielles continues sur D et la condition 2) équivaut à ce que le jacobien de l'application f, i.e. le déterminant det $f'(w)$, ne s'annule nulle part sur D.

Il résulte des conditions 1) et 2) de la définition 1 en vertu du théorème 9.3.2 que f est localement injective, i.e. tout point $w \in D$ possède un voisinage $U \subset D$ tel que la restriction de f à U est injective. Cependant, f n'est pas obligatoirement injective, comme le montre l'exemple d'une application f: $\mathbf{R}^2 \to \mathbf{R}^2$ de composantes

$$f_1(r, \varphi) = r\cos\varphi, \qquad f_2(r, \varphi) = r\sin\varphi,$$

qui a été envisagé dans la remarque au théorème 9.3.4. Ainsi donc, la condition 3) ne découle pas de 1) et 2).

2. Propriétés principales de l'application régulière

THÉORÈME 1. *Soit f une application de $\mathbf{R}^n$ dans $\mathbf{R}^n$ définie et régulière sur un ensemble ouvert non vide D. Alors* :

1) *l'image de tout ensemble ouvert contenu dans D est un ensemble ouvert* ;

2) *l'application réciproque de f est aussi régulière* ;

3) *l'image de la frontière de tout compact contenu dans D sert de frontière à l'image de ce compact et l'image de l'intérieur sert d'intérieur.*

Démonstration. 1) La restriction $f|_O$ de l'application f à tout ouvert $O \subset D$ et tout point $w \in O$ vérifient les hypothèses du théorème 9.3.4 : puisque $f|_O$ est différentiable au point w et $(f|_O)'(w) = f'(w)$ (voir exemple 7.2.4), on obtient conformément à la condition 1 de la définition 1 que $(f|_O)'$ est continue au point w et que l'opérateur $(f|_O)'(w)$ est inversible d'après la condition 2. Par conséquent, il découle du théorème 9.3.4 que tout point $w \in O$ possède un voisinage ouvert $U_w \subset O$ tel que $f(U_w) = (f|_O)(U_w)$ est un ensemble ouvert. Or, vu que $O = \bigcup_{w \in O} U_w$, on a $f(O) = \bigcup_{w \in O} f(U_w)$. Par suite, $f(O)$ (et, en particulier, $f(D)$) est un ensemble ouvert.

2) Conformément à la condition 3) de la définition 1, f est injective, de sorte qu'il existe une application g réciproque de f ; à noter que d'après les assertions précédentes, $D_g = f(D)$ est un ensemble ouvert. Quel que soit le point $w \in D$, l'application g_w réciproque de la restriction de f à un voisinage U_w envisagé plus haut de ce point, est une restriction de g à $f(U_w)$. Or, le théorème 9.3.4 dit que g_w est différentiable et que $f(U_w)$ est un voisinage du point $z = f(w)$. Par conséquent, g est différentiable au point z et, d'après le même théorème, g' est continue en ce point. Puisque w est un point quelconque de D et que z est un point quelconque de $f(D)$, on en conclut que g est continûment différentiable sur D_g, i.e. g satisfait à la condition 1) de la définition 1. Comme $g'(z) = [f'(g(z))]^{-1}$ d'après le théorème 9.3.4, $g'(z)$ est inversible pour chacun des points $z \in D_g$: $[g'(z)]^{-1} = f'(g(z))$. Ainsi donc, g satisfait aussi à la condition 2). Enfin, g est injective en tant qu'application réciproque de f, i.e. elle satisfait aussi à la condition 3). Donc, g est régulière.

3) Soit x un compact contenu dans D. Puisque $f(]x[) \subset f(x)$ et que $f(]x[)$ est un ouvert en tant qu'image de l'ensemble ouvert $]x[$ conformément à ce qui a été démontré en 1), on a $f(]x[) \subset]f(x)[$. Vu que f est continue, $f(x)$ est un compact d'après le théorème 3.2.6 et, par suite, c'est un ensemble fermé. C'est pourquoi, $\mathrm{Fr}\, f(x) \subset f(x)$. Or $x = \mathrm{Fr}\, x \cup]x[$. Donc,

$$\mathrm{Fr}\, f(x) \subset f(\mathrm{Fr}\, x) \cup f(]x[) \subset f(\mathrm{Fr}\, x) \cup]f(x)[.$$

Puisque $\mathrm{Fr}\, f(x)$ n'a pas de points communs avec $]f(x)[$, on en conclut que

$$\mathrm{Fr}\, f(x) \subset f(\mathrm{Fr}\, x). \tag{1}$$

D'autre part, $x = g(f(x))$. Comme $f(x)$ est un compact et g est régulière d'après 2), on obtient conformément à ce qui vient d'être démontré (que

l'on applique à $f(x)$ au lieu de x et à g au lieu de f) :

$$\text{Fr } x = \text{Fr } g(f(x)) \subset g(\text{Fr } f(x)),$$

d'où

$$f(\text{Fr } x) \subset f(g(\text{Fr } f(x))) = \text{Fr } f(x). \tag{2}$$

En comparant (1) et (2), on conclut que $f(\text{Fr } x) = \text{Fr } f(x)$. Enfin, vu que

$$f(\text{Fr } x) \cup f(]x[) = f(\text{Fr } x \cup]x[) = f(x) = \text{Fr } f(x) \cup]f(x)[$$

et que de plus $f(\text{Fr } x) = \text{Fr } f(x)$, $f(]x[) \subset]f(x)[$ et $\text{Fr } f(x)$ n'a pas de points communs avec $f(]x[)$, on obtient $f(]x[) =]f(x)[$.

REMARQUE. Si les hypothèses du théorème 1 sont satisfaites, *l'image $f(O)$ de chaque domaine $O \subset D$ est un domaine.* En effet, d'après le théorème, $f(O)$ et O sont des ouverts. Comme O est par ailleurs connexe, il en est de même de $f(O)$ d'après le théorème 3.1.1 puisque f est continue. Donc, $f(O)$ est un domaine. En particulier, si D est un domaine, $f(D)$ est aussi un domaine.

§ 2. Variation de l'aire par une application affine

1. Aire du parallélogramme

Soient Π un plan de coordonnées rapporté à un repère direct (resp. rétrograde) et P un parallélogramme dans Π engendré par un couple de vecteurs $r_1 = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$, $r_2 = \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$ ayant une origine commune (fig. 40). Comme tout polygone, P est quarrable. On donne à son aire le signe « + » si (r_1, r_2) est un couple direct (resp. rétrograde) et le signe « − » si le couple est rétrograde (resp. direct). Alors

$$\text{aire } P = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}. \tag{1}$$

En effet, aire $P = \|r_1\| \cdot \|r_2\| \cdot \sin\alpha = \|r_1\| \cdot \|r_2\| \cdot \cos\left(\alpha - \frac{\pi}{2}\right)$ où α est un angle auquel il faut tourner r_1 pour que les vecteurs r_1 et r_2 soient parallèles de même sens. Vu que l'aire ne varie pas lors d'une translation, on peut considérer sans restreindre la généralité que r_1 et r_2 sont issus de l'origine des coordonnées O. Soit T une rotation du plan Π autour de O à l'angle $-\pi/2$. On a alors $\|Tr_2\| = \|r_2\|$ et $\widehat{(r_1, Tr_2)} = \alpha - \frac{\pi}{2}$, si bien que l'aire de

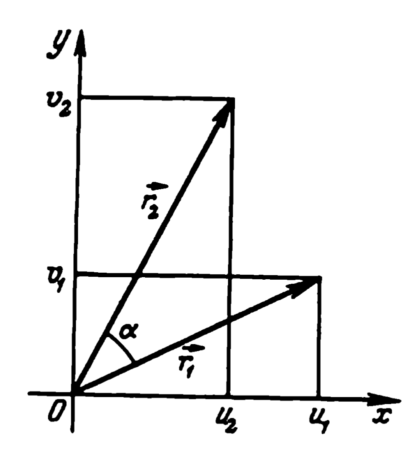


Fig. 40

P est égale au produit scalaire des vecteurs r_1 et Tr_2 :

$$\text{aire}\, P = \|r_1\| \cdot \|Tr_2\| \cdot \cos(\widehat{r_1,\ Tr_2}) = \langle r_1,\ Tr_2 \rangle.$$

Or

$$Tr_2 = \begin{pmatrix} \cos\left(-\frac{\pi}{2}\right) & -\sin\left(-\frac{\pi}{2}\right) \\ \sin\left(-\frac{\pi}{2}\right) & \cos\left(-\frac{\pi}{2}\right) \end{pmatrix} \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ -u_2 \end{pmatrix}.$$

Par conséquent,

$$\text{aire}\, P = \left\langle \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \begin{pmatrix} v_2 \\ -u_2 \end{pmatrix} \right\rangle = u_1 v_2 - u_2 v_1 = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}.$$

2. Variation de l'aire du parallélogramme par une application affine

Une application affine A du plan de coordonnées Π dans le plan de coordonnées Π_1 est définie par les formules

$$(2) \qquad \begin{cases} x = a_{11}u + a_{12}v + a_1, \\ y = a_{21}u + a_{22}v + a_2, \end{cases}$$

où u, v sont les coordonnées d'un point du plan Π, x, y, les coordonnées de son image dans Π_1 et a_{ij} et a_i sont des constantes. Ceci étant, on a

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Les plans Π et Π_1 étant munis de coordonnées, on peut les identifier à $\mathbf{R}^2$. Il s'ensuit alors que A (envisagée en tant qu'application de $\mathbf{R}^2$ dans

$\mathbf{R}^2$) est continûment différentiable, et $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ est sa matrice jacobienne car

$$a_{11} = x'_u, \quad a_{12} = x'_v, \quad a_{21} = y'_u, \quad a_{22} = y'_v.$$

Ainsi donc, $\det A$ est le jacobien de l'application A. Comme $\det A \neq 0$, A est une application régulière.

On sait du cours de géométrie que l'application affine envoie les segments en des segments et les vecteurs colinéaires en des vecteurs colinéaires. Il en résulte en particulier que l'image du parallélogramme est un parallélogramme et que si l'un de deux parallélogrammes est le translaté de l'autre, il en est de même de leurs images.

THÉORÈME 1. *Une application affine du plan fait multiplier l'aire de tout parallélogramme par son jacobien, i.e. si* $A : \Pi \to \Pi_1$ *est une application affine définie par les formules* (2) *et* P *est un parallélogramme dans* Π (*de sorte que* $A(P)$ *est un parallélogramme dans* Π_1), *on a*

$$\text{aire}\, A(P) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \text{aire}\, P = \det A \cdot \text{aire}\, P.$$

D é m o n s t r a t i o n. Sans restreindre la généralité, on peut admettre que P est engendré par les vecteurs $r_1 = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$, $r_2 = \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$ issus de l'origine des coordonnées O. L'application A fait associer à ces vecteurs les vecteurs $A(r_i)$ ($i = 1, 2$) issus de $O'(a_1, a_2)$ dont les coordonnées sont

$$x_i - a_1 = a_{11}u_i + a_{12}v_i,$$

$$y_i - a_2 = a_{21}u_i + a_{22}v_i.$$

Ceci étant, $A(P)$ est un parallélogramme engendré par les vecteurs $A(r_1)$, $A(r_2)$. En prenant en considération la formule (1) et la règle de multiplication des déterminants, on obtient

$$\text{aire}\, A(P) = \begin{vmatrix} x_1 - a_1 & x_2 - a_1 \\ y_1 - a_2 & y_2 - a_2 \end{vmatrix} = \begin{vmatrix} a_{11}u_1 + a_{12}v_1 & a_{11}u_2 + a_{12}v_2 \\ a_{21}u_1 + a_{22}v_1 & a_{21}u_2 + a_{22}v_2 \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} = \det A \cdot \text{aire}\, P.$$

Il résulte du théorème 1 que $|\text{aire}\, A(P)| = |\det A| \cdot |\text{aire}\, P|$. Par la suite, on admettra que les vecteurs qui engendrent le parallélogramme ne sont pas ordonnés et que l'aire du parallélogramme est donc toujours égale à sa valeur absolue.

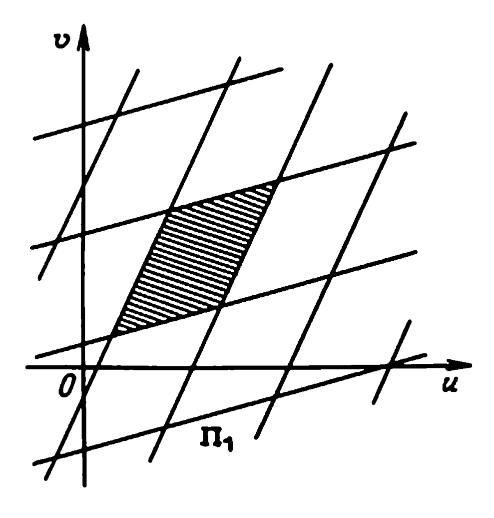


Fig. 41

3. Variation de l'aire de la figure quarrable par une application affine

THÉORÈME 2. *L'image de la figure quarrable par une application affine est une figure quarrable. Si A est une application affine du plan Π dans le plan Π_1, et Φ est une figure quarrable, on a*

$$\text{aire}\, A(\Phi) = |\det A| \cdot \text{aire}\, \Phi.$$

Démonstration. Subdivisons le plan Π en des carrés

$$Q_{k,l} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \times \left[\frac{l}{2^n}, \frac{l+1}{2^n}\right] \quad (k,\ l \in \mathbf{Z},\ n \in \mathbf{N})$$

de rang n. Leurs images $A(Q_{k,l})$ sont des parallélogrammes sans points intérieurs communs, qui s'obtiennent les uns des autres par translation et dont la réunion recouvre Π_1 (fig. 41). Rappelons que $A_n(\Phi)$, où Φ est une figure dans Π, signifie la réunion de tous les carrés de rang n qui sont contenus à l'intérieur de Φ, et $B_n(\Phi)$ signifie celle des carrés qui ont des points communs avec Φ. Comme $A_n(\Phi) \subset \Phi \subset B_n(\Phi)$, on a

$$A(A_n(\Phi)) \subset A(\Phi) \subset A(B_n(\Phi)).$$

En tant que figures composées d'un nombre fini de parallélogrammes, $A(A_n(\Phi))$ et $A(B_n(\Phi))$ sont quarrables. On a donc en vertu du théorème 1 et de l'additivité de l'aire :

$$\text{aire}\, A(A_n(\Phi)) = |\det A| \cdot \text{aire}\, A_n(\Phi),$$

$$\text{aire}\, A(B_n(\Phi)) = |\det A| \cdot \text{aire}\, B_n(\Phi).$$

Si Φ est quarrable, il existe pour tout $\varepsilon > 0$ un $n \in \mathbf{N}$ tel que

$$\text{aire}\, B_n(\Phi) - \text{aire}\, A_n(\Phi) < \frac{\varepsilon}{|\det A|}$$

et, par conséquent, $\text{aire}\, A(B_n(\Phi)) - \text{aire}\, A\ (A_n(\Phi)) < \varepsilon$. Ainsi donc, la figure $A(\Phi)$ peut être encadrée par deux figures quarrables dont les aires diffèrent aussi peu que l'on veut, et par suite, elle est quarrable. On a dans ce cas :

$$\text{aire}\, A(\Phi) = \lim \text{aire}\, A(B_n(\Phi)) = |\det A| \cdot \lim \text{aire}\, B_n(\Phi) = |\det A| \cdot \text{aire}\, \Phi.$$

§ 3. Invariance de la quarrabilité et variation de l'aire par une application régulière

1. Changement de normes des vecteurs et des applications linéaires

Pour estimer la variation de l'aire par une application régulière, il s'avère rationnel de modifier la définition de la norme d'un vecteur et, respectivement, celle de la norme d'une application linéaire. Plus précisément, on appellera *norme d'un vecteur* $x \in \mathbf{R}^n$ (en gardant l'ancienne notation) l'expression

$$\|x\| = \max_{i \in [1,n]} |x_i|$$

où x_i sont les coordonnées du vecteur x dans la base standard. Il est aisé de voir que c'est justement la norme. La « boule » ouverte de rayon h et de centre x^0 sera alors représentée par un cube $Q(x^0 ; h)$ de dimension n. En effet,

$$Q(x^0 ; h) = \{x \in \mathbf{R}^n |\, |x_1 - x_1^0| < h, \ldots, |x_n - x_n^0| < h\} =$$
$$= \{x \in \mathbf{R}^n |\, \max_{i \in [1,n]} |x_i - x_i^0| < h\} = \{x \in \mathbf{R}^n |\, \|x - x^0\| < h\}.$$

Soient A une application linéaire de $\mathbf{R}^n$ dans $\mathbf{R}^m$ et $(a_{ij})i\in[1,m], j\in[1,n]$ sa matrice dans le couple de bases standard, de sorte que Ax est un vecteur (dans $\mathbf{R}^m$) de coordonnées $\sum_{j=1}^{n} a_{ij}x_j$ $(i = 1, \ldots, m)$ pour tout $x \in \mathbf{R}^n$. Les espaces $\mathbf{R}^n$ et $\mathbf{R}^m$ étant munis de la nouvelle norme, posons

$$P_A = \{\varrho \geqslant 0 |\, \|Ax\| \leqslant \varrho \|x\| \text{ pour tout } x \in \mathbf{R}^n\}$$

et

(1) $$\|A\| = \inf P_A.$$

Comme

$$\|Ax\| = \max_{i\in[1,m]}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \leqslant \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|\cdot|x_j| \leqslant$$

$$\leqslant \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|\cdot\max_{j\in[1,n]}|x_j| = \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|\cdot\|x\|,$$

on a

(2) $$\max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}| \in P_A,$$

de sorte que $P_A \neq \varnothing$. Tout comme dans 7.1.3, il en résulte que $\|A\|$ est une norme sur $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$, $\|Ax\| \geqslant \|A\|\cdot\|x\|$ pour tout $x\in\mathbf{R}^n$, $\|BA\| \leqslant \leqslant \|B\|\cdot\|A\|$ pour toutes les applications $A\in\mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ et $B\in\mathscr{L}(\mathbf{R}^m, \mathbf{R}^l)$, et $\|I_{\mathbf{R}^n}\| = 1$. En vertu de (1) et (2),

(3) $$\|A\| \leqslant \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|.$$

Montrons qu'en réalité c'est une égalité. En cas de $A = 0$, c'est évident. Soit $A \neq 0$ et $\sum_{j=1}^{n}|a_{i_0j}| = \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|$, de sorte que $a_{i_0j_0} \neq 0$ pour un entier $j_0 \in [1, n]$. Posons $x_j = \operatorname{sgn} a_{i_0j}$ $(j = 1, \ldots, n)$ et $x = \sum_{j=1}^{n} x_je^j$. Alors $\|x\| = |\operatorname{sgn} a_{i_0j_0}| = 1$ et, en vertu de (3),

$$\|A\| \leqslant \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}| = \sum_{j=1}^{n}|a_{i_0j}| = \left|\sum_{j=1}^{n} a_{i_0j}x_j\right| \leqslant \|Ax\| \leqslant \|A\|\cdot\|x\| = \|A\|,$$

d'où

(4) $$\|A\| = \max_{i\in[1,m]}\sum_{j=1}^{n}|a_{ij}|.$$

La métrique engendrée dans $\mathbf{R}^n$ par la nouvelle norme des vecteurs est liée à la métrique euclidienne par les inégalités suivantes :

$$\|x - x^0\| = \max_{i\in[1,n]}|x_i - x_i^0| \leqslant \sqrt{\sum_{i=1}^{n}(x_i - x_i^0)^2} \leqslant \sqrt{n}\cdot\|x - x^0\|.$$

C'est pourquoi, lorsque x tend vers x^0 pour l'une de ces métriques, il en est de même pour l'autre. En particulier, $\|x\|$ *est une fonction continue* (dans le sens usuel) *des coordonnées* $x_1, \ldots, x_n$ *du vecteur* x. Il en résulte que $\|A\|$ *est une fonction continue des éléments* a_{ij} *de la matrice de l'application* A. En effet, on voit de la formule (4) que $\|A\|$ est une fonction

composée obtenue de $\max\limits_{i\in[1,m]} |y_i|$ (qui dépend continûment de $y_1, ..., y_m$) si on y remplace tout y_i par la somme correspondante $\sum\limits_{j=1}^{n} |a_{ij}|$ (qui dépend continûment de $a_{i1}, ..., a_{in}$). Il est évident que

$$[Q(x^0\,;\,h)] = \{x \in \mathbf{R}^n \mid \|x - x^0\| \leqslant h\}.$$

2. Invariance de la quarrabilité

Par la suite, dans ce paragraphe, f désignera partout une application du plan de coordonnées Π dans le plan de coordonnées Π_1, *définie et régulière sur un ensemble ouvert D*; les points du plan Π seront désignés par la lettre w (munie parfois des indices) et leurs coordonnées, par les lettres u, v; les composantes de l'application f seront notées f_1 et f_2, et la matrice jacobienne de f sera désignée par $\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}$ (de sorte que $f_{11}(w) = \frac{\partial f_1}{\partial u}(w)$ et $f_{12}(w) = \frac{\partial f_1}{\partial v}(w)$ et de façon analogue pour f_{21} et f_{22}).

LEMME 1. *Si $C = [Q(w_0\,;\,h)] \subset D$, on a $f(C) \subset [Q(f(w_0)\,;\,h')]$ où* $h' = \max \|f'(w')\| \cdot h$, *de sorte que*

$$\text{aire}\,[Q(f(w_0)\,;\,h')] = \left(\max_{w'\in C} \|f'(w')\|\right)^2 \text{aire}\, C.$$

Démonstration. Soit $w \in C$, i.e. $\|w - w_0\| \leqslant h$, de sorte que le segment $[w_0, w]$ est contenu dans C. Conformément à la formule de Lagrange (voir théorème 9.2.1),

$$f_i(w) - f_i(w_0) = f_{i1}(w_i)(u - u_0) + f_{i2}(w_i)(v - v_0) \quad (i = 1,\ 2)$$

où w_1, w_2 sont des points intérieurs du segment $[w_0, w]$. C'est pourquoi,

$$|f_i(w) - f_i(w_0)| \leqslant (|f_{i1}(w_i)| + |f_{i2}(w_i)|) \max\{|u - u_0|, |v - v_0|\} =$$
$$= (|f_{i1}(w_i)| + |f_{i2}(w_i)|)\|w - w_0\|$$

et, par conséquent,

$$\|f(w) - f(w_0)\| = \max\{|f_1(w) - f_1(w_0)|, |f_2(w) - f_2(w_0)|\} \leqslant$$
$$\leqslant \max\{|f_{11}(w_1)| + |f_{12}(w_1)|, |f_{21}(w_2)| + |f_{22}(w_2)|\}\|w - w_0\|.$$

Comme $w_i \in C$, on a

$$|f_{i1}(w_i)| + |f_{i2}(w_i)| \leqslant \max\{|f_{i1}(w')| + |f_{i2}(w')|\} \leqslant$$
$$\leqslant \max_{w'\in C} \max\{|f_{11}(w')| + |f_{12}(w')|, |f_{21}(w')| + |f_{22}(w')|\} =$$
$$= \max_{w'\in C} \|f'(w')\| \quad (i = 1,\ 2)$$

(où les valeurs maximales sur C existent puisque f_{ij} sont continues et C est un compact), donc

$$\max\{|f_{11}(w_1)| + |f_{12}(w_1)|, |f_{21}(w_2)| + |f_{22}(w_2)|\} \leqslant \max_{w' \in C} \|f'(w')\|.$$

Par suite,

$$\|f(w) - f(w_0)\| \leqslant \max_{w' \in C} \|f'(w')\| \cdot \|w - w_0\|.$$

Si maintenant $z \in f(C)$, i.e. $z = f(w)$ où $w \in C$, on obtient

$$\|z - f(w_0)\| \leqslant \max \|f'(w')\| \cdot \|w - w_0\| \leqslant \max_{w' \in C} \|f'(w')\| \cdot h = h',$$

i.e. $z \in [Q(f(w_0) ; h')]$.

THÉORÈME 1. *L'image $f(x)$ de tout compact quarrable $x \subset D$ est un compact quarrable.*

Démonstration. Pour $x = \varnothing$ l'assertion est triviale. Soit $x \neq \varnothing$. On entend par $\varrho(A, B)$, $A, B \subset \Pi$, la distance de A à B pour la métrique *euclidienne*. Comme $x \subset D$ et D est un ensemble ouvert, on a $x \cap \operatorname{Fr} D = \varnothing$. Puisque x est un compact et $\operatorname{Fr} D$ est un ensemble fermé, il en résulte d'après la propriété 3.2.5, 5° que $\varrho = \varrho(x, \operatorname{Fr} D) > 0$. Prenons un $\delta \in]0, \varrho[$ arbitraire et soit

$$x' = B(x ; \delta) = \{w \in \Pi \mid \varrho(w, x) \leqslant \delta\}.$$

Vu que $x \subset D$ et $\delta < \varrho$, le théorème 3.3.4 dit que $x' \subset U(x ; \varrho) \subset D$. Ceci étant, x' est un compact d'après la propriété 3.2.5, 3°. Comme f_{ij} sont des fonctions continues, on obtient conformément à ce qu'il a été dit à la fin du n° 1 que $\|f'(w)\|$ est aussi une fonction continue et, par conséquent, elle présente sur x' la valeur maximale. L'opérateur $f'(w)$ étant non nul partout sur D parce qu'inversible, $\max\limits_{w \in x} \|f'(w)\| \neq 0$. Puisque x est quarrable, il existe pour tout $\varepsilon > 0$ en vertu du théorème 11.1.2 un $n \in \mathbf{N}$ tel que

$$\text{aire } B_n(\operatorname{Fr} x) < \varepsilon / \left(\max_{w \in x'} \|f'(w)\|\right)^2. \tag{5}$$

Ceci étant, on peut prendre n aussi grand que les diagonales des carrés de rang n soient strictement inférieures à δ et alors $B_n(\operatorname{Fr} x) \subset x'$ (car tout point $w \in B_n(\operatorname{Fr} x)$ appartient à un carré de rang n ayant des points communs avec $\operatorname{Fr} x$ et, par conséquent, la distance de w à x est strictement inférieure à δ). Soient $C_1, \ldots, C_m$ des carrés de rang n formant $B_n(\operatorname{Fr} x)$ et soient $w_1, \ldots, w_m$ leurs centres, de sorte que $C_i = [Q(w_i ; 2^{-n-1})]$. D'après le lemme 1, $f(C_i) \subset [Q(f(w_i) ; h_i')]$ où

$$h_i' = \max_{w' \in C_i} \|f'(w')\| \cdot 2^{-n-1}.$$

Comme $\operatorname{Fr} f(x) = f(\operatorname{Fr} x)$ d'après le théorème 1 du § 1, on a

$$\operatorname{Fr} f(x) \subset f(B_n(\operatorname{Fr} x)) = \bigcup_{i=1}^{m} f(C_i) \subset \bigcup_{i=1}^{m} [Q(f(w_i)\ ;\ h_i'].$$

Vu que $B_n(\operatorname{Fr} x) \subset x'$,

$$\text{aire} \bigcup_{i=1}^{m} [Q(f(w_i)\ ;\ h_i')] \leqslant \sum_{i=1}^{m} \text{aire}\,[Q(f(w_i)\ ;\ h_i')] =$$

$$= \sum_{i=1}^{m} \left(\max_{w' \in C_i} \|f'(w')\|\right)^2 \cdot \text{aire}\, C_i \leqslant \left(\max_{w \in x'} \|f'(w)\|\right)^2 \cdot \sum_{i=1}^{m} \text{aire}\, C_i =$$

$$= \left(\max_{w \in x'} \|f'(w)\|\right)^2 \text{aire}\, B_n(\operatorname{Fr} x) < \varepsilon.$$

Ainsi donc, la frontière de $f(x)$ peut être recouverte par une figure quarrable $\bigcup_{i=1}^{m} [Q(f(w_i)\ ;\ h_i')]$ dont l'aire est aussi petite que l'on veut, i.e. elle est quarrable et son aire est nulle. En vertu du théorème 11.1.2, cela signifie que $f(x)$ est une figure quarrable. Enfin, comme f est continue, $f(x)$ est un compact d'après le théorème 3.2.6.

3. Estimation de l'aire d'une image

Il résulte en particulier du théorème 1 que l'image $f(C)$ de tout carré C de la forme $[Q(w_0\ ;\ h)]$ est quarrable. On a de plus en vertu du lemme 1,

$$\text{(6)} \qquad \text{aire}\, f(C) \leqslant \left(\max_{w \in C} \|f'(w)\|\right)^2 \cdot \text{aire}\, C.$$

Cette estimation admet la généralisation suivante.

LEMME 2. *Tout carré $C \subset \Pi$ de la forme* $[Q(w_0\ ;\ h)]$ *et toute transformation linéaire non dégénérée A du plan* Π *vérifient l'inégalité*

$$\text{(7)} \qquad \text{aire}\, f(C) \leqslant |\det A| \cdot \left(\max_{w \in C} \|A^{-1} f'(w)\|\right)^2 \cdot \text{aire}\, C.$$

D é m o n s t r a t i o n. Comme A^{-1} est une transformation affine, le théorème 2 du § 2 affirme que la figure $A^{-1}(\Phi)$ est quarrable si Φ est une figure quarrable. Ceci étant, on a

$$\text{aire}\, A^{-1}(\Phi) = |\det A^{-1}| \cdot \text{aire}\, \Phi.$$

Puisque $\Phi = f(C)$ est une figure quarrable, on obtient

$$\text{(8)} \qquad \text{aire}\, f(C) = |\det A| \cdot \text{aire}\, A^{-1}(f(C)).$$

Mais $A^{-1} f$ et f sont définies et régulières sur D, si bien qu'en appliquant

l'inégalité (6) à $A^{-1}f$, on peut écrire

$$\text{(9)}\quad \text{aire}\, A^{-1}(f(C)) \leqslant \left(\max_{w\in C} \|(A^{-1}f)'(w)\|\right)^2 \cdot \text{aire}\, C = \\ = \left(\max_{w\in C} \|A^{-1}f'(w)\|\right)^2 \cdot \text{aire}\, C.$$

Enfin, (8) et (9) entraînent (7).

LEMME 3. *Tout compact quarrable $x \subset D$ vérifie l'inégalité*

$$\text{(10)}\qquad \text{aire}\, f(x) \leqslant \iint_x |\det f'|.$$

Démonstration. Supposons que δ et x' soient définis comme dans la démonstration du théorème 1. En se donnant le nombre $\varepsilon > 0$, on y prend un $n \in \mathbf{N}$ tel que les diamètres de tous les carrés de rang n soient strictement inférieurs à δ (d'où il résulte que $B_n(\text{Fr}\, x) \subset x'$) et l'inégalité (5) soit satisfaite. On y établit que

$$f(B_n(\text{Fr}\, x)) \subset \bigcup_{i=1}^{m} [Q(f(w_i')\,;\, h_i')] \text{ où aire} \bigcup_{i=1}^{m} [Q(f(w_i')\,;\, h_i')] < \varepsilon.$$

Vu que $f(B_n(\text{Fr}\, x))$ est quarrable en vertu du théorème 1, il en résulte que

$$\text{(11)}\qquad \text{aire}\, f(B_n(\text{Fr}\, x)) < \varepsilon.$$

On a vu dans la démonstration du théorème 11.1.2 que les carrés de rang n qui ont des points communs avec x et qui ne sont pas contenus dans $]x[$, se coupent avec $\text{Fr}\, x$. C'est pourquoi,

$$B_n(x) \subset A_n(x) \cup B_n(\text{Fr}\, x),$$

d'où

$$f(B_n(x)) \subset f(A_n(x)) \cup f(B_n(\text{Fr}\, x)).$$

Puisque les trois figures sont quarrables, on a d'après la remarque au théorème 11.1.3,

$$\text{aire}\, f(B_n(x)) \leqslant \text{aire}\, f(A_n(x)) + \text{aire}\, f(B_n(\text{Fr}\, x))$$

et, par suite, en vertu de (11)

$$\text{(12)}\qquad \text{aire}\, f(B_n(x)) - \text{aire}\, f(A_n(x)) < \varepsilon.$$

Si aire $x = 0$, on a $A_n(x) = \varnothing$, d'où aire $f(A_n(x)) =$ aire $\varnothing = 0$ et l'inégalité (12) devient aire $f(B_n(x)) < \varepsilon$. Comme $x \subset B_n(x)$, on obtient $(0 \leqslant)$ aire $f(x) < \varepsilon$ et, vu que cela est juste pour tout $\varepsilon > 0$, aire $f(x) = 0$, si bien que l'inégalité (10) (même égalité stricte) est démontrée. Supposons maintenant que aire $x > 0$. Comme aire $x = \sup_n$ aire $A_n(x)$, $A_n(x)$ n'est pas vide pour un n suffisamment grand. Soient $C_1, \ldots, C_p$ des carrés de rang n qui forment $A_n(x)$. Comme ils sont connexes et $|\det f'|$ est continu, le théo-

rème de la valeur moyenne (voir propriété 11.4.3, 10°) dit qu'il existe des points $w_i \in C_i$ tels que

$$\iint_{C_i} |\det f'| = |\det f'(w_i)| \cdot \text{aire } C_i.$$

En posant $C = C_i$ et $A = f'(w_i)$ dans l'inégalité (7), on obtient

$$\text{(13)} \quad \text{aire} f(C_i) \leqslant |\det f'(w_i)| \left(\max_{w \in C_i} \|[f'(w_i)]^{-1} f'(w)\|\right)^2 \cdot \text{aire } C_i =$$

$$= \left(\max_{w \in C_i} \|[f'(w_i)]^{-1} f'(w)\|\right)^2 \cdot \iint_{C_i} |\det f'|.$$

Estimons $\|[f'(w_i)]^{-1} f'(w)\|^2 - 1$ où $w \in C_i$. Etant des fonctions continues, $\|f'(w)\|$ et $\|f'(w)]^{-1}\|$ présentent sur x leurs valeurs maximales, par exemple c et d. Vu que $C_i \subset A_n(x) \subset x$ et en prenant en considération que la norme de la composée d'applications linéaires ne dépasse pas le produit de leurs normes et que la norme de l'opérateur identique est égale à 1 (voir n° 7.1.3, propriétés 5° et 6°), on a

$$\text{(14)} \quad |\|[f'(w_i)]^{-1} f'(w)\|^2 - 1| =$$

$$= (\|[f'(w_i)]^{-1} f'(w)\| + 1) |\|[f'(w_i)]^{-1} f'(w)\| - 1| \leqslant$$

$$\leqslant (\|[f'(w_i)]^{-1}\| \cdot \|f'(w)\| + 1) |\|[f'(w_i)]^{-1} f'(w)\| - \|[f'(w_i)]^{-1} f'(w_i)\|| \leqslant$$

$$\leqslant (cd + 1) \cdot \|[f'(w_i)]^{-1} f'(w) - [f'(w_i)]^{-1} f'(w_i)\| \leqslant$$

$$\leqslant (cd + 1) \cdot d \|f'(w) - f'(w_i)\|.$$

Etant continu, l'opérateur $f'(w)$ est uniformément continu sur le compact x, de sorte que δ peut être considéré aussi petit que l'on ait

$$\text{(15)} \quad (w, \ w' \in x \text{ et } \|w - w'\| < \delta) \Rightarrow$$

$$\Rightarrow \|f'(w) - f'(w')\| < \frac{\varepsilon}{(cd + 1) d \iint_x |\det f'|}$$

(le second membre de l'inégalité a le sens car $d \neq 0$ puisque $[f'(w)]^{-1} \neq 0$, et $\iint_x |\det f'| \neq 0$ d'après la propriété 11.4.3, 8'° puisque $|\det f'| > 0$ et est continu et que $]x[\neq \varnothing$). Comme w et w_i sont contenus dans C_i et $C_i \subset x$, la prémisse de l'implication (15) est vraie et c'est pourquoi il résulte de (14) que

$$|\|[f'(w_i)]^{-1} f'(w)\|^2 - 1| < \frac{\varepsilon}{\iint_x |\det f'|} \quad \text{pour tout } w \in C_i,$$

d'où

$$\left(\max_{w \in C_i} \|[f'(w_i)]^{-1} f'(w)\|\right)^2 < 1 + \frac{\varepsilon}{\iint_x |\det f'|}.$$

En substituant cette inégalité dans (13), on obtient

$$\operatorname{aire} f(C_i) < \left(1 + \frac{\varepsilon}{\iint_x |\det f'|}\right) \cdot \iint_{C_i} |\det f'| \qquad (i = 1, \dots, p)$$

et, par conséquent (puisque $A_n(x) \subset x$),

$$(16) \quad \operatorname{aire} f(A_n(x)) = \sum_{i=1}^{p} \operatorname{aire} f(C_i) <$$

$$< \left(1 + \frac{\varepsilon}{\iint_x |\det f'|}\right) \iint_{A_n(x)} |\det f'| \leqslant \iint_x |\det f'| + \varepsilon.$$

En additionnant (12) et (16) et en prenant en considération que $x \subset B_n(x)$, on obtient

$$\operatorname{aire} f(x) \leqslant \operatorname{aire} f(B_n(x)) < \iint_x |\det f'| + 2\varepsilon,$$

d'où il résulte (10) en vertu de ce que ε est arbitraire.

4. Aire de l'image. Coefficient de déformation des aires

LEMME 4. *Si F est une fonction réelle bornée et positive sur un compact quarrable $\sigma \subset f(D)$, on a*

$$(17) \qquad \iint_\sigma F = \iint_{f^{-1}(\sigma)} (F \circ f) |\det f'|.$$

Démonstration. Notons avant tout que d'après le théorème 1 du § 1, $g = f^{-1}$ est régulière si f est régulière, de sorte que $x = f^{-1}(\sigma)$ est un compact quarrable d'après le théorème 1. La fonction F étant bornée sur σ, il en est de même pour $(F \circ f)|\det f'|$ sur x en vertu de la continuité de $\det f'$, si bien que le second membre de la formule (17) a le sens. Soit T une partition quelconque du compact σ en $\sigma_1, \dots, \sigma_n$. Soient $m_i = \inf_{\sigma_i} F$ et $g(\sigma_i) = x_i$ (de sorte que $\sigma_i = f(x_i)$, x_i sont des compacts quarrables et $\bigcup_{i=1}^{n} x_i = x$). Comme $F \geqslant 0$, on a $m_i \geqslant 0$ et, en vertu du lemme 3,

$$(18) \quad \underline{s}_T^{(F)} = \sum_{i=1}^{n} m_i \operatorname{aire} \sigma_i = \sum_{i=1}^{n} m_i \operatorname{aire} f(x_i) \leqslant$$

$$\leqslant \sum_{i=1}^{n} m_i \iint_{x_i} |\det f'| = \sum_{i=1}^{n} \iint_{x_i} m_i |\det f'|.$$

Vu que $m_i = \inf\limits_{x_i} (F \circ f)$, on a $m_i |\det f'| \leqslant (F \circ f)|\det f'|$ sur x_i, ce qui implique

$$\iint_{x_i} m_i |\det f'| = \underline{\iint}_{x_i} m_i |\det f'| \leqslant \underline{\iint}_{x_i} (F \circ f)|\det f'|.$$

Or, si $i \neq j$, on obtient en vertu du théorème 1 du § 1 que

$$]x_i[\cap]x_j[=]g(\sigma_i)[\cap]g(\sigma_j)[= g(]\sigma_i[) \cap g(]\sigma_j[) = g(]\sigma_i[\cap]\sigma_j[) = g(\varnothing) = \varnothing,$$

de sorte que x_i et x_j n'ont pas de points intérieurs communs. Par suite, d'après la propriété 11.4.2 (III),

$$\underline{s}_T^{(F)} \leqslant \sum_{i=1}^{n} \underline{\iint}_{x_i} (F \circ f)|\det f'| = \underline{\iint}_{x} (F \circ f)|\det f'|$$

et donc on a aussi

(19) $$\underline{\iint}_{\sigma} F = \sup_T \underline{s}_T^{(F)} \leqslant \underline{\iint}_{x} (F \circ f)|\det f'|.$$

Puisque g est régulière, l'inégalité analogue est vérifiée par toute fonction réelle bornée positive G sur x :

(19′) $$\underline{\iint}_{x} G \leqslant \underline{\iint}_{\sigma} (G \circ g)|\det g'|.$$

Si $G = (F \circ f)|\det f'|$, on a $(G \circ g)|\det g'| = F$. En effet (en prenant en considération la formule de dérivation de la composée d'applications et le fait que les fonctions f et g sont réciproques l'une de l'autre et que le déterminant de la composée d'applications linéaires est égal au produit de leurs déterminants), on a pour tout $z \in f(D)$

$$G(g(z))|\det g'(z)| = F(f(g(z)))|\det f'(g(z))|\,|\det g'(z)| =$$
$$= F(z)|\det(f'(g(z))g'(z))| = F(z)|\det(f \circ g)'(z)| =$$
$$= F(z)|\det I'_{\Pi_1}(z)| = F(z)|\det I_{\Pi_1}| = F(z).$$

Ainsi donc, il résulte, en particulier de (19), que

(20) $$\underline{\iint}_{x} (F \circ f)|\det f'| \leqslant \underline{\iint}_{\sigma} F.$$

Les inégalités (19) et (20) donnent l'égalité (17).

Maintenant, on peut renforcer sensiblement le lemme 3.

THÉORÈME 2. *Tout compact quarrable $x \subset D$ vérifie l'égalité*

(21) $$\text{aire} f(x) = \iint_{x} |\det f'|.$$

Démonstration. En posant $\sigma = f(x)$ et $F \equiv 1$ dans (17) (de sorte que $F \circ f \equiv 1$), on obtient

$$\operatorname{aire} f(x) = \operatorname{aire} \sigma = \iint_\sigma 1 = \underline{\iint}_\sigma 1 = \underline{\iint}_x (1 \circ f)|\det f'| =$$

$$= \underline{\iint}_x |\det f'| = \iint_x |\det f'|.$$

EXEMPLE 1. Soit à calculer l'aire de la figure σ limitée par les hyperboles $xy = 1$, $xy = 2$ et les paraboles $y^2 = x$, $y^2 = 2x$. Cette figure est située dans le premier quadrant ouvert du plan (x, y) (voir fig. 42). Considérons une application g du premier quadrant dans le plan (u, v), ayant des composantes données par les formules

$$u = xy, \qquad v = \frac{y^2}{x}.$$

Elle est continûment différentiable et

$$\det g'(x, y) = \frac{D(u, v)}{D(x, y)} = \begin{vmatrix} y & x \\ -\dfrac{y^2}{x^2} & \dfrac{2y}{x} \end{vmatrix} = \frac{3y^2}{x} = 3v \neq 0.$$

En outre, g est injective puisque $x^3 = u^2/v$ et $y^3 = uv$. Ainsi donc, g et, par conséquent, l'application f qui lui est inverse, sont régulières et $\det f'(u, v) = \dfrac{1}{\det g'(x, y)} = \dfrac{1}{3v}$. On a $\sigma = f(x)$ où x est un carré du plan (u, v) limité par les droites $u = 1$, $u = 2$, $v = 1$, $v = 2$. Par conséquent, la figure σ est quarrable selon le théorème 1 et

$$\operatorname{aire} \sigma = \iint_x \frac{1}{3v} = \int_1^2 \int_1^2 \frac{dv}{3v}\, du = \frac{1}{3} \ln 2$$

d'après le théorème 2.

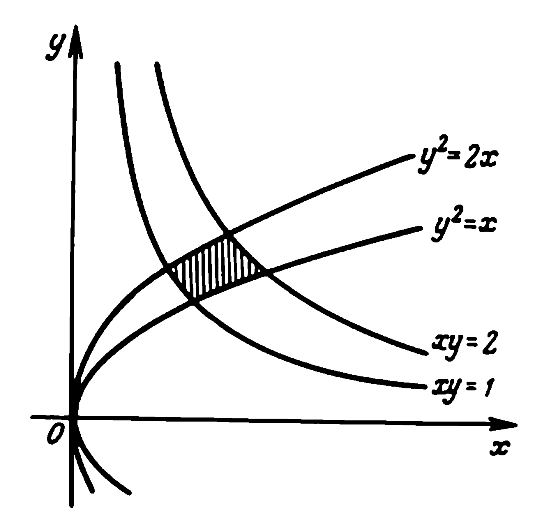


Fig. 42

COROLLAIRE. *On a* $\dfrac{\text{aire} f(x)}{\text{aire } x} \to |\det f'(w_0)|$ *quand* x *se réduit en un point* $w_0 \in D$, *i.e. pour tout compact quarrable* $x \subset D$
$(\forall \varepsilon > 0)(\exists \delta > 0)(\text{aire } x \neq 0,\ w_0 \in x$ *et* $\text{diam } x < \delta) \Rightarrow$

$$\Rightarrow \left| \frac{\text{aire} f(x)}{\text{aire } x} - |\det f'(w_0)| \right| \leqslant \varepsilon.$$

En effet, comme $|\det f'|$ est continu au point w_0, on a
$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall w \in D)(d(w, w_0) < \delta) \Rightarrow$

$$\Rightarrow |\det f'(w_0)| - \varepsilon \leqslant |\det f'(w)| \leqslant |\det f'(w_0)| + \varepsilon.$$

Si aire $x \neq 0$, $w_0 \in x$ et diam $x < \delta$, de sorte que $d(w, w_0) < \delta$ pour tout $w \in x$, on obtient d'après la propriété 11.4.3, 9° que

$$|\det f'(w_0)| - \varepsilon \leqslant \frac{1}{\text{aire } x} \iint_x |\det f'(w_0)| + \varepsilon$$

et, par conséquent, en vertu du théorème 2,

$$\left| \frac{\text{aire} f(x)}{\text{aire } x} - |\det f'(w_0)| \right| = \left| \frac{1}{\text{aire } x} \iint_x |\det f'| - |\det f'(w_0)| \right| \leqslant \varepsilon.$$

REMARQUE. Si f est une application affine, on obtient en vertu du théorème 2 du § 2 que $\dfrac{\text{aire} f(x)}{\text{aire } x} = |\det f'(w_0)|$. D'après ce qui vient d'être démontré, cette égalité n'est remplie que d'une façon approximative pour une application régulière f arbitraire. Plus diam x est petit, et plus cette approximation est précise. Le nombre $|\det f'(w_0)|$ est appelé *coefficient de déformation de l'aire* au point w_0 par l'application régulière f.

§ 4. Formules de changement des variables

1. Changement des variables par une application régulière

LEMME 1. *Toute fonction* F *réelle bornée sur un compact quarrable* $\sigma \subset f(D)$ *vérifie les égalités*

$$\text{(1)} \qquad \underline{\iint}_\sigma F = \underline{\iint}_{f^{-1}(\sigma)} (F \circ f)|\det f'|,$$

$$\text{(2)} \qquad \overline{\iint}_\sigma F = \overline{\iint}_{f^{-1}(\sigma)} (F \circ f)|\det f'|.$$

Démonstration. La condition de F positive n'a été nécessaire dans la démonstration du lemme 4 du § 3 que pour obtenir le signe $\leqslant$ dans la formule (18). Or le théorème 2 du § 3 affirme que la formule (18) est vraie indépendamment de cette condition même si on y remplace $\leqslant$ par $=$. Puisque aire $f(x_i) = \iint_{x_i} |\det f'|$, on a

$$\underline{s}_T^{(F)} = \sum_{i=1}^{n} m_i \operatorname{aire} f(x_i) = \sum_{i=1}^{n} \iint_{x_i} m_i |\det f'|$$

pour toute F bornée. Le résultat obtenu permet de poursuivre les raisonnements donnés dans la démonstration du lemme 4 du § 3 pour établir la validité de l'égalité (1) pour toutes les fonctions F bornées. Quant à l'égalité (2), on l'obtient de (1) en vertu de la propriété 11.4.1 (II) par substitution de $-F$ à F :

$$\overline{\iint_\sigma} F = -\underline{\iint_\sigma}(-F) = -\underline{\iint}_{f^{-1}(\sigma)}(-F\circ f)|\det f'| = \overline{\iint}_{f^{-1}(\sigma)}(F\circ f)|\det f'|.$$

Le théorème suivant résulte directement du lemme 1.

THÉORÈME 1. *Soit f une application du plan Π dans le plan Π_1, qui est définie et régulière sur un ensemble ouvert non vide D et soit σ un compact quarrable contenu dans $f(D)$. Une fonction F est intégrable sur σ si et seulement si la fonction $G = (F\circ f)|\det f'|$ est intégrable sur $f^{-1}(\sigma)$. Ceci étant, on a alors*

$$\iint_\sigma F = \iint_{f^{-1}(\sigma)} G = \iint_{f^{-1}(\sigma)} (F\circ f)|\det f'|. \tag{3}$$

2. Applications quasi régulières

La formule (3) reste en vigueur même si l'on affaiblit un tout petit peu la condition imposée à f d'être une application régulière.

DÉFINITION 1. Soient Π et Π_1 des plans de coordonnées (qui peuvent par là même être identifiés à $\mathbf{R}^2$). Une application f de Π dans Π_1 sera dite *quasi régulière sur l'adhérence* $[D]$ d'un ouvert non vide D si : 1) f est continûment différentiable sur un voisinage D' de l'ensemble $[D]$; 2) $\det f'(w) \neq 0$ sur D ; 3) f est injective sur D et 4) $x \cap \operatorname{Fr} D$ est une figure d'aire nulle pour tout compact quarrable $x \subset [D]$.

Il est évident que f est alors régulière sur D.

EXEMPLE 1. Soient Π et Π_1, deux plans rapportés respectivement aux coordonnées cartésiennes r, φ et x, y, et soit

$$D = \{(r, \varphi) \in \Pi \mid r > 0,\ 0 < \varphi < 2\pi\},$$

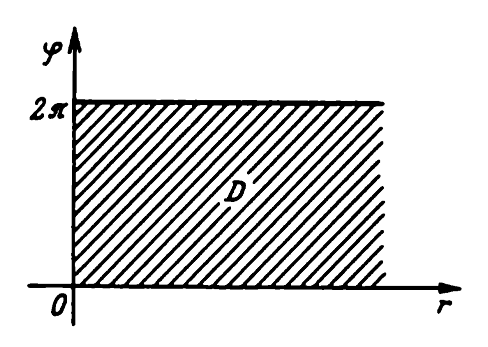


Fig. 43

de sorte que $[D] = \{(r, \varphi) \in \Pi \mid r \geqslant 0, 0 \leqslant \varphi \leqslant 2\pi\}$ (fig. 43). L'application $f: \Pi \to \Pi_1$ de composantes $x = r\cos\varphi$, $y = r\sin\varphi$ est continûment différentiable et son jacobien $\det f'(r, \varphi) = r \neq 0$ sur D. Si $(r_i, \varphi_i) \in D$ ($i = 1, 2$) et $f(r_1, \varphi_1) = f(r_2, \varphi_2)$, i.e.

$$(r_1\cos\varphi_1,\ r_1\sin\varphi_1) = (r_2\cos\varphi_2,\ r_2\sin\varphi_2),$$

on a

$$r_1^2 = (r_1\cos\varphi_1)^2 + (r_1\sin\varphi_1)^2 = (r_2\cos\varphi_2)^2 + (r_2\sin\varphi_2)^2 = r_2^2,$$

d'où $r_1 = r_2$ (puisque $r_1 r_2 > 0$) et alors

$$(\cos\varphi_1,\ \sin\varphi_1) = (\cos\varphi_2,\ \sin\varphi_2),$$

d'où il résulte que φ_1 et φ_2 se distinguent par un multiple pair de π et, comme $|\varphi_1 - \varphi_2| < 2\pi$, on a $\varphi_1 = \varphi_2$. C'est pourquoi, f est injective sur D. Enfin, $\operatorname{Fr} D$ comprend le segment $\{(0, \varphi) \mid 0 \leqslant \varphi \leqslant 2\pi\}$ de l'axe φ et les demi-droites $\{(r, 0) \mid r > 0\}$ et $\{(r, 2\pi) \mid r > 0)\}$, de sorte que la condition 4 de la définition 1 est aussi vérifiée. Ainsi, f est quasi régulière sur $[D]$, et $f([D]) = \Pi_1$. Notons que f n'est pas injective sur $[D]$ car $(f(r, -\pi) \equiv f(r, \pi)$ et $f(0, \varphi) \equiv 0$.

3. Invariance de la quarrabilité par une application quasi régulière

Dans les n°s 3 et 4, Π et Π_1 désigneront des plans de coordonnées, et D, un ensemble ouvert non vide dans Π.

THÉORÈME 2. *Si une application $f: \Pi \to \Pi_1$ est quasi régulière sur $[D]$, l'image $f(x)$ de tout compact quarrable x contenu dans $[D]$ est un compact quarrable.*

Démonstration. Soit x un compact quarrable contenu dans $[D]$ et soit $U([D]\,;\,\varrho)$ un voisinage D' de l'ensemble $[D]$ sur lequel f est continûment différentiable. Vu que f est continue sur x, $f(x)$ est un compact. Si $x \cap \operatorname{Fr} D = \varnothing$, i.e. $x \subset D$, $f(x)$ est un compact quarrable d'après le théorème 1 du § 3 puisque f est régulière sur D. Soit $x \cap \operatorname{Fr} D \neq \varnothing$, de sorte

que $B_n(x \cap \mathrm{Fr}\, D) \neq \emptyset$ pour tout $n \in \mathbf{N}$. Posons $x' = B(x, \delta)$ où $\delta \in]0, \varrho[$; conformément à la propriété 3.2.5, 3°, x' est un compact, et comme $\delta < \varrho$, on a $x' \subset D'$. A son tour, $B_n(x \cap \mathrm{Fr}\, D) \subset x'$ pour tout n suffisamment grand tel que les diagonales des carrés de rang n soient inférieures à δ (car tout point $w \in B_n(x \cap \mathrm{Fr}\, D)$ appartient à un carré de rang n ayant des points communs avec x et est alors éloigné de x à une distance $< \delta$). Soit $w_0 \in x \cap \mathrm{Fr}\, D$. Comme $x \cap \mathrm{Fr}\, D \subset]B_n(x \cap \mathrm{Fr}\, D)[$, on obtient que $B_n(x \cap \cap \mathrm{Fr}\, D)$ et, à plus forte raison, x' sont des voisinages du point $w_0 \in \mathrm{Fr}\, D$, de sorte que x' a une intersection non vide avec D. Par suite, $\|f'(w)\| \not\equiv 0$ sur x' en vertu de la condition 2 de la définition 1, et c'est pourquoi $\max_{w \in x'} \|f'(w)\|$ (qui existe en vertu de la continuité de $\|f'(w)\|$) est strictement positif. Etant donné que aire $(x \cap \mathrm{Fr}\, D) = 0$, on obtient pour chaque $\varepsilon > 0$ et pour tout $n \in \mathbf{N}$ suffisamment grand :

$$\text{(4)} \qquad \text{aire}\, B_n(x \cap \mathrm{Fr}\, D) < \varepsilon / \left(\max_{w \in x'} \|f'(w)\|\right)^2.$$

Soit n_ε le plus petit de tous les $n \in \mathbf{N}$ qui vérifient cette inégalité et l'inclusion $B_n(x \cap \mathrm{Fr}\, D) \subset x'$. Vu que $x \subset D'$ et f est continûment différentiable sur D', le lemme 1 du § 3 (fondé sur la seule différentiabilité continue de f) implique

$$f(B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)) \subset \bigcup_{i=1}^{m} [Q(f(w_i)\, ;\, h_i')],$$

où w_i sont les centres des carrés C_i de rang n formant $B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)$, et $h_i' = \max_{w \in C_i} \|f'(w)\| \cdot 2^{-n-1}$. Comme $B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D) \subset x'$, on a alors

$$\overline{\text{aire}}\, f(B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)) \leqslant \sum_{i=1}^{m} \text{aire}\, [Q(f(w_i)\, ;\, h_i')] =$$

$$= \sum_{i=1}^{m} \left(\max_{w \in C_i} \|f'(w)\|\right)^2 \text{aire}\, C_i \leqslant \left(\max_{w \in x'} \|f'(w)\|\right)^2 \sum_{i=1}^{m} \text{aire}\, C_i =$$

$$= \left(\max_{w \in x'} \|f'(w)\|\right)^2 \text{aire}\, B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D) < \varepsilon.$$

Posons maintenant $x_\varepsilon = x \setminus]B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)[$. En tant que sous-ensemble fermé d'un compact, x_ε est un compact et, en tant que différence des figures quarrables, c'est une figure quarrable. On a de plus $x_\varepsilon \subset x \setminus (x \cap \cap \mathrm{Fr}\, D) \subset D$. Par conséquent, $f(x_\varepsilon)$ est un compact quarrable d'après le théorème 1 du § 3. Vu que $x_\varepsilon \subset x \subset x_\varepsilon \cup B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)$ on obtient $f(x_\varepsilon) \subset f(x) \subset f(x_\varepsilon) \cup f(B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D))$ et c'est pourquoi

$$\text{(5)} \quad \text{aire}\, f(x_\varepsilon) \leqslant \underline{\text{aire}}\, f(x) \leqslant \overline{\text{aire}}\, f(x) \leqslant$$

$$\leqslant \text{aire}\, f(x_\varepsilon) + \overline{\text{aire}}\, f(B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)) < \text{aire}\, f(x_\varepsilon) + \varepsilon,$$

d'où $(0 \leqslant)\ \overline{\text{aire}}\, f(x) - \underline{\text{aire}}\, f(x) < \varepsilon$. Puisque ε est arbitraire, on en conclut que $\overline{\text{aire}}\, f(x) = \underline{\text{aire}}\, f(x)$, de sorte que $f(x)$ est un compact quarrable.

4. Changement des variables par une application quasi régulière

LEMME 1′. *Supposons que les hypothèses du théorème* 2 *soient satisfaites et que* x *soit un compact quarrable dans* $[D]$, *de sorte que* $f(x)$ *est un compact quarrable d'après le théorème* 2. *Toute fonction* F *réelle bornée sur* $f(x)$ *vérifie alors les égalités*

$$(1')\qquad \underline{\iint}_{f(x)} F = \underline{\iint}_{x} (F \circ f)|\det f'|,$$

$$(2')\qquad \overline{\iint}_{f(x)} F = \overline{\iint}_{x} (F \circ f)|\det f'|.$$

Démonstration. Gardons les notations introduites pour la démonstration du théorème 2. En vertu du théorème 11.1.4 et des inégalités (5), on a

$$\text{aire}\,[f(x) \setminus f(x_\varepsilon)] = \text{aire}\,(f(x) \setminus f(x_\varepsilon)) = \text{aire}\, f(x) - \text{aire}\, f(x_\varepsilon) < \varepsilon.$$

D'autre part, il existe par hypothèse un $C > 0$ tel que $|F| < C$ sur $f(x)$. Par conséquent, en vertu des propriétés 4° et 11° de l'intégrale double,

$$\left|\underline{\iint}_{f(x)} F - \underline{\iint}_{f(x_\varepsilon)} F\right| = \left|\underline{\iint}_{[f(x) \setminus f(x_\varepsilon)]} F\right| \leqslant C_\varepsilon,$$

d'où il résulte que

$$(6)\qquad \underline{\iint}_{f(x)} F = \lim_{\varepsilon \to 0} \underline{\iint}_{f(x_\varepsilon)} F.$$

De façon analogue, en utilisant l'inclusion $x \setminus x_\varepsilon \subset B_{n_\varepsilon}(x \cap \mathrm{Fr}\, D)$, ainsi que l'inégalité (4) pour $n = n_\varepsilon$ et le fait que la fonction $(F \circ f)|\det f'|$ est bornée sur x, on démontre que

$$(7)\qquad \underline{\iint}_{x} (F \circ f)|\det f'| = \lim_{\varepsilon \to 0} \underline{\iint}_{x_\varepsilon} (F \circ f)|\det f'|.$$

Comme $x_\varepsilon \subset D$ et f est régulière sur D, on obtient d'après le lemme 1 que

$$(8)\qquad \underline{\iint}_{f(x_\varepsilon)} F = \underline{\iint}_{x_\varepsilon} (F \circ f)|\det f'|.$$

L'égalité (1′) résulte des formules (6) à (8). Enfin, tout comme à la fin de la démonstration du lemme 1, on obtient (2′) de (1′) par substitution de $-F$ à F.

Le théorème suivant découle directement du lemme 1′.

THÉORÈME 1′. *Soit $f: \Pi \to \Pi_1$ une application quasi régulière sur l'adhérence $[D]$ d'un ensemble ouvert non vide D et soit x un compact quarrable dans $[D]$, de sorte que $f(x)$ est un compact quarrable d'après le théorème 2. Une fonction F est intégrable sur $f(x)$ si et seulement si la fonction $(F \circ f)|\det f'|$ est intégrable sur x. Ceci étant, on a*

$$\iint_{f(k)} F = \iint_{x} (F \circ f)|\det f'|. \tag{9}$$

EXERCICE 1. Soit f une application de Π dans Π_1 définie sur un ensemble ouvert non vide D et satisfaisant aux conditions suivantes : 1) f est continûment différentiable et $\neq 0$, 2) l'intersection de l'ensemble

$$K = \{w \in D|\ \det f'(x) = 0\}$$

avec tout compact quarrable $x \subset D$ est une figure d'aire nulle et 3) f est injective sur $D \setminus K$. Démontrer que l'image $f(x)$ de tout compact quarrable $x \subset D$ est un compact quarrable et que la fonction F est intégrable sur $f(x)$ si et seulement si la fonction $(F \circ f)|\det f'|$ est intégrable sur x. Montrer qu'on a dans ce cas la formule (9).

5. Passage aux coordonnées polaires dans l'intégrale double

Soient Π et Π_1 deux plans rapportés respectivement aux coordonnées cartésiennes r, φ et x, y et soit

$$D = \{(r, \varphi) \in \Pi|\ r > 0,\ 0 < \varphi < 2\pi\}.$$

On a vu dans l'exemple 1 que l'application $f: \Pi \to \Pi_1$ de composantes $x = r\cos\varphi$, $y = r\sin\varphi$ est quasi régulière sur $[D]$ et que $\det f'(r, \varphi) = r$. Pour cette raison, il résulte du théorème 1′ que tout compact quarrable $x \subset [D]$ et toute fonction F intégrable sur $f(x)$ vérifient la formule

$$\iint_{f(x)} F(x, y)\,dx\,dy = \iint_{x} F(r\cos\varphi,\ r\sin\varphi)\,r\,dr\,d\varphi. \tag{10}$$

EXEMPLE 2. Soit à calculer le volume V d'un corps découpé dans la boule

$$x^2 + y^2 + z^2 \leqslant a^2 \text{ (où } a > 0)$$

par la surface $x^2 + y^2 - ax = 0$, i.e. par le cylindre circulaire droit

$$\left(x - \frac{a}{2}\right)^2 + y^2 = \left(\frac{a}{2}\right)^2.$$

Puisque y et z ne sont que du deuxième degré, le corps envisagé est symétrique par rapport aux plans de coordonnées $y = 0$ et $z = 0$, de sorte que le problème se ramène au calcul du volume $V/4$ d'une portion $\mathscr{T}$ du corps, qui est située dans le premier octant fermé $\{(x, y, z)|\ x \geqslant 0,\ y \geqslant 0,\ z \geqslant 0\}$ (fig. 44). Vu que $\mathscr{T}$ est le sous-graphe de la fonction $F(x, y) = \sqrt{a^2 - x^2 - y^2}$ sur un demi-disque σ du plan (x, y) donné par les inéquations $x^2 + y^2 - ax \leqslant 0$, $y \geqslant 0$, on obtient d'après le corollaire du théorème 11.3.1 que $\mathscr{T}$ est en effet cubable et que

$$V/4 = \operatorname{vol}\mathscr{T} = \iint_\sigma \sqrt{a^2 - x^2 - y^2}\,dx\,dy.$$

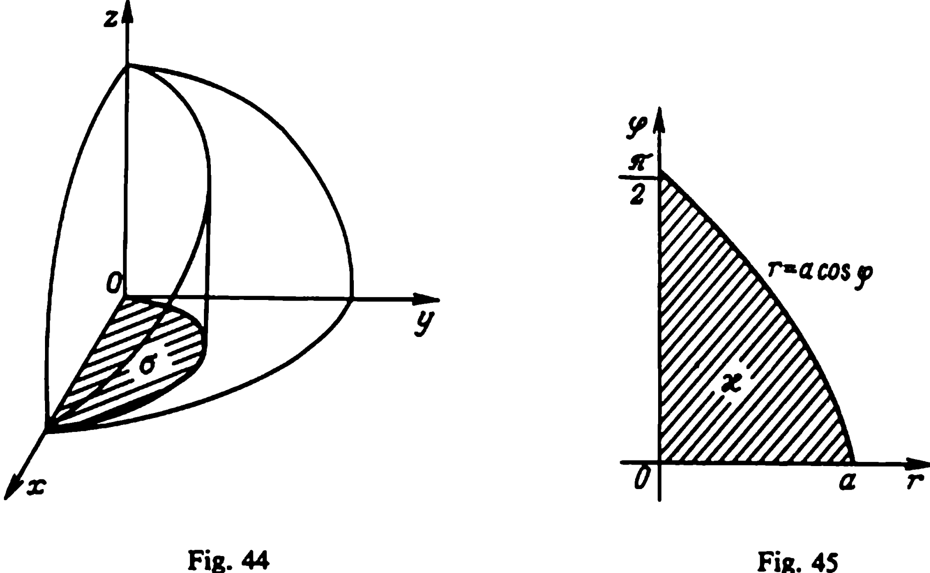


Fig. 44 Fig. 45

Mais σ est une image du sous-graphe x (du côté de l'axe φ) de la restriction de la fonction $a \cos \varphi$ au segment $[0, \pi/2]$ (fig. 45) par une application f envisagée plus haut. Comme f est quasi régulière sur l'ensemble

$$[D] = \{(r, \varphi) \in \Pi | \ r \geqslant 0, \ 0 \leqslant \varphi \leqslant 2\pi\}$$

(exemple 1) et x est un compact quarrable contenu dans $[D]$, on a d'après la formule (10) :

$$V/4 = \iint_x \sqrt{a^2 - r^2}\, r \, \mathrm{d}r \, \mathrm{d}\varphi = \int_0^{\frac{\pi}{2}} \int_0^{a \cos \varphi} \sqrt{a^2 - r^2}\, r \, \mathrm{d}r \, \mathrm{d}\varphi.$$

En transformant l'intégrale intérieure par la substitution $u = \sqrt{a^2 - r^2}$, on obtient

$$\int_0^{a \cos \varphi} \sqrt{a^2 - r^2}\, r \, \mathrm{d}r = \int_{a \sin \varphi}^{0} u^2 \, \mathrm{d}u = \frac{a^3}{3}(1 - \sin^3 \varphi).$$

Par conséquent,

$$V = \frac{4a^3}{3} \int_0^{\pi/2} (1 - \sin^3 \varphi) \mathrm{d}\varphi = \frac{4a^3}{3}\left(\frac{\pi}{2} - \frac{2}{3}\right).$$

EXERCICE 2. Montrer que les propriétés figurant dans l'exercice 1 sont vérifiées pour l'application f de l'exemple 1 et pour l'ensemble

$$D = \{(r, \varphi) \in \Pi | \ \alpha < \varphi < \alpha + \pi\},$$

α étant quelconque. Utiliser ce résultat pour résoudre le problème envisagé dans l'exemple 2.

CHAPITRE 13

INTÉGRALE TRIPLE

§ 1. Notion d'intégrale triple

1. Intégrales inférieure et supérieure. Intégrabilité

La notion d'intégrale double donnée dans le chapitre 11 est fondée sur la notion de compact quarrable. De façon analogue, la notion d'intégrale triple sera fondée sur la notion de compact cubable.

On appelle *partition d'un compact cubable* v toute représentation de ce compact sous la forme de la réunion d'une famille finie $(v_k)_{k\in[1,n]}$ de compacts cubables dont les intérieurs sont disjoints deux à deux. Les compacts v_k seront appelés *éléments de la partition*, le volume de v_k sera noté Δv_k. Le plus grand diamètre d_T des éléments de la partition T est appelé *diamètre* de cette partition *). L'ensemble $\mathcal{T}_v$ de toutes les partitions du compact cubable v possède toutes les propriétés de l'ensemble $\mathcal{T}_\sigma$ des partitions du compact quarrable σ, qui ont été établies au n° 11.3.1.

Soit f une fonction bornée arbitraire sur un compact cubable v et soit T une partition de v en $v_1, \ldots, v_n$. Les sommes

$$\underline{s}_T = \sum_{k=1}^{n} m_k \Delta v_k, \qquad \bar{s}_T = \sum_{k=1}^{n} M_k \Delta v_k,$$

où $m_k = \inf\limits_{v_k} f$, $M_k = \sup\limits_{v_k} f$ s'appellent respectivement *sommes inférieure* et *supérieure* de f correspondant à T. Tout comme dans le cas bidimensionnel, on démontre que $\underline{s}_{T_1} \leqslant \bar{s}_{T_2}$ pour toutes partitions T_1 et T_2. Il s'ensuit que :

1) la borne supérieure de la famille $(\underline{s}_T)_{T\in\mathcal{T}_v}$ de toutes les sommes inférieures est finie. On l'appelle *intégrale inférieure* de la fonction f sur v et on la note $\underline{\iiint_v} f$;

2) la borne inférieure de la famille $(\bar{s}_T)_{T\in\mathcal{T}_v}$ de toutes les sommes supérieures est finie. On l'appelle *intégrale supérieure* de la fonction f sur v

*) La notion de diamètre d'un ensemble de tout espace métrique est donnée au n° 11.3.1.

et on la note $\overline{\iiint_v} f$;

3) $\underline{\iiint_v} f \leqslant \overline{\iiint_v} f$.

Si $\underline{\iiint_v} f = \overline{\iiint_v} f$, on dit que f est *intégrable* sur v et que la valeur commune de ses intégrales inférieure et supérieure est l'*intégrale triple* de la fonction f sur v ; elle est notée

$$\iiint_v f(x, y, z)\,dx\,dy\,dz, \text{ ou } \iiint_v f\,dv, \text{ ou encore } \iiint_v f.$$

L'intégrale triple possède les propriétés analogues à celles de l'intégrale double décrites aux §§ 3 et 4 du chapitre 11 et qu'on démontre de la même façon (à condition bien sûr de remplacer partout un élément quarrable par un élément cubable et, d'une façon plus générale, une figure plane par un corps). En particulier, *toute fonction continue sur un compact cubable est intégrable.*

2. Intégrale triple de Riemann

Soient f une fonction bornée sur un compact cubable v et T une partition de v en $v_1, \ldots, v_n$. La somme

$$s_{T,\Xi} = \sum_{k=1}^{n} f(\xi_k)\Delta v_k,$$

où Ξ est un ensemble de points ξ_k choisis arbitrairement dans chaque v_k, est appelée *somme de Riemann* de f attachée à la partition (T, Ξ). On dit que f est *intégrable au sens de Riemann* sur v si $s_{T,\Xi}$ tend vers un nombre I lorsque $d_T \to 0$, i.e.

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall (T, \Xi)(d_T < \delta) \Rightarrow |s_{T,\Xi} - I| < \varepsilon).$$

Le nombre I est alors appelé *intégrale triple de Riemann* de la fonction f sur v. Tout comme en dimensions un et deux, on démontre que f est intégrable au sens de Riemann si et seulement si ses intégrales inférieure et supérieure coïncident sur v et que l'intégrale triple de Riemann de la fonction f sur v est alors égale à leur valeur commune.

§ 2. Calcul de l'intégrale triple par intégration répétée

Le calcul des intégrales triples pour certains domaines d'intégration et fonctions se ramène au calcul successif d'intégrales simples et doubles.

1. Calcul de l'intégrale triple sur un corps cylindrique

THÉORÈME 1. *Soit σ un compact quarrable dans le plan xOy et soit $[c_1, c_2]$ un segment de l'axe z. Si une fonction $f(x, y, z)$ est intégrable sur le corps cylindrique $v = \sigma \times [c_1, c_2]$ et si pour tout point donné $(x, y) \in \sigma$ elle est intégrable sur le segment $[c_1, c_2]$ en tant que fonction de z, la fonction*

$$g(x, y) = \int_{c_2}^{c_1} f(x, y, z)\mathrm{d}z$$

est intégrable sur σ et

$$\iiint_\sigma f = \iint_\sigma g = \iint_\sigma \left(\int_{c_1}^{c_2} f(x, y, z)\,\mathrm{d}z\right) \mathrm{d}x\,\mathrm{d}y.$$

Démonstration. Elle est analogue dans l'ensemble à celle du théorème 11.5.1. Soient T_1 une partition de σ en σ_i $(i \in [1, m])$ et T_2 une subdivision du segment $[c_1, c_2]$ par des points z_j $(j \in [0, n])$, de sorte que

$$c_1 = z_0 < z_1 < \ldots < z_{n-1} < z_n = c_2.$$

Posons

$$v_{ij} = \sigma_i \times [z_j, z_{j+1}] \quad (i \in [1, m],\ j \in [0, n-1])$$

(fig. 46). D'après le théorème 11.2.1, les corps cylindriques v_{ij} sont cubables et $\Delta v_{ij} = (z_{j+1} - z_j)\Delta\sigma_i$. Leur compacité résulte aisément de la compacité

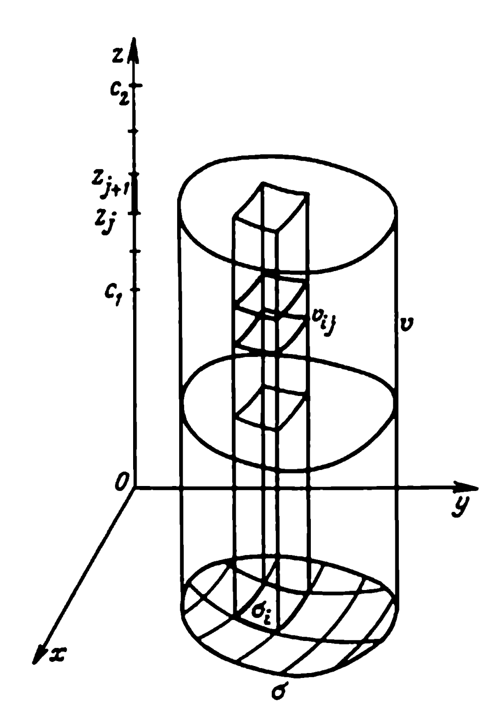


Fig. 46

des ensembles σ_i et $[z_j, z_{j+1}]$. Il est aussi évident qu'ils n'ont pas de points intérieurs communs et que leur réunion est égale à v. Ainsi donc, v_{ij} sont les éléments d'une partition T_{12} du corps v. Soient enfin

$$m_{ij} = \operatorname*{int}_{v_{ij}} f, \qquad M_{ij} = \sup_{v_{ij}} f, \qquad m_i = \inf_{\sigma_i} g, \qquad M_i = \sup_{\sigma_i} g.$$

Vu que $m_{ij} \leqslant f(x, y, z) \leqslant M_{ij}$ pour tout $(x, y, z) \in v_{ij}$, on a

$$m_{ij}\Delta z_j \leqslant \int_{z_j}^{z_{j+1}} f(x, y, z)\,\mathrm{d}z \leqslant M_{ij}\Delta z_j$$

pour tout $(x, y) \in \sigma_i$. C'est pourquoi,

$$\sum_{j=0}^{n+1} m_{ij}\Delta z_j \leqslant \sum_{j=0}^{n+1} \int_{z_j}^{z_{j+1}} f(x, y, z)\,\mathrm{d}z = g(x, y) \leqslant \sum_{j=0}^{n-1} M_{ij}\Delta z_j$$

pour tout $(x, y) \in \sigma_i$, d'où il résulte que

$$\sum_{j=0}^{n-1} m_{ij}\Delta z_j \leqslant m_i \leqslant M_i \leqslant \sum_{j=0}^{n-1} M_{ij}\Delta z_j.$$

En multipliant chaque membre de cette triple inégalité par $\Delta\sigma_i$ et en faisant la sommation par rapport à i, on obtient

$$(1) \qquad \underline{s}^{(f)}_{T_{12}} = \sum_{i=1}^{m} \sum_{j=0}^{n-1} m_{ij}\Delta z_j \Delta\sigma_i \leqslant \underline{s}^{(g)}_{T_1} \leqslant \overline{s}^{(g)}_{T_1} \leqslant \sum_{i=1}^{m} \sum_{j=0}^{n-1} M_{ij}\Delta z_j \Delta\sigma_i = \overline{s}^{(f)}_{T_{12}}.$$

Vu que f est intégrable sur v, on a

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall T \in \mathscr{T}_v)(d_T < \delta \Rightarrow \overline{s}^{(f)}_T - \underline{s}^{(f)}_T < \varepsilon).$$

Or $d_{T_{12}} = \max\sqrt{(\operatorname{diam}\sigma_i)^2 + \Delta z_j^2} < \delta$ si $d_{T_1} < \delta/\sqrt{2}$ et $d_{T_2} < \delta/\sqrt{2}$. Par conséquent,

$$\overline{s}^{(f)}_{T_{12}} - \underline{s}^{(f)}_{T_{12}} < \varepsilon$$

et donc, en vertu de (1), on obtient à plus forte raison :

$$\overline{s}^{(g)}_{T_1} - \underline{s}^{(g)}_{T_1} < \varepsilon.$$

Puisque ε est arbitraire, cela signifie que g est intégrable sur σ, si bien qu'on a à la suite de (1) :

$$\underline{s}^{(f)}_{T_{12}} \leqslant \iint_\sigma g \leqslant \overline{s}^{(f)}_{T_{12}}.$$

La double inégalité

$$\underline{s}^{(f)}_{T_{12}} \leqslant \iiint_v f \leqslant \overline{s}^{(f)}_{T_{12}}$$

étant aussi vraie, on en conclut que $\left|\iiint_v f - \iint_\sigma g\right| < \varepsilon$ pour tout $\varepsilon > 0$, i.e.

$$\iiint_v f = \iint_\sigma g.$$

THÉORÈME 1′. *Soient σ un compact quarrable dans le plan xOy et $[c_1, c_2]$ un segment de l'axe z. Si une fonction $f(x, y, z)$ est intégrable sur le corps cylindrique $v = \sigma \times [c_1, c_2]$ et si pour tout $z \in [c_1, c_2]$ fixé elle est intégrable sur σ en tant que fonction du point (x, y), la fonction*

$$h(z) = \iint_\sigma f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y$$

est intégrable sur le segment $[c_1, c_2]$ et

$$\iiint_v f = \int_{c_1}^{c_2} \left(\iint_\sigma f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y \right) \mathrm{d}z.$$

Démonstration. Elle est analogue à la démonstration du théorème 1. On a maintenant

$$m_j = \inf_{[z_j,\, z_{j+1}]} h\,, \qquad M_j = \sup_{[z_j, z_{j+1}]} h \qquad (j = 0, 1, \ldots, n-1),$$

les autres notations restent les mêmes. Comme $m_{ij} \leqslant f(x, y, z) \leqslant M_{ij}$ pour tout $(x, y, z) \in v_{ij}$, on a

$$m_{ij}\Delta\sigma_i \leqslant \iint_{\sigma_i} f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y \leqslant M_{ij}\sigma_i$$

pour tout $z \in [z_j, z_{j+1}]$, d'où

$$\sum_{i=1}^{m} m_{ij}\Delta\sigma_i \leqslant h(z) \leqslant \sum_{i=1}^{m} M_{ij}\Delta\sigma_i$$

pour tout $z \in [z_j, z_{j+1}]$ et, par conséquent,

$$\sum_{i=1}^{m} m_{ij}\Delta\sigma_i \leqslant m_j \leqslant M_j \leqslant \sum_{i=1}^{m} M_{ij}\Delta\sigma_i.$$

En multipliant chaque membre de cette triple inégalité par Δz_j et en faisant la sommation par rapport à j, on obtient

$$\underline{s}_{T_{12}}^{(f)} \leqslant \underline{s}_{T_2}^{(h)} \leqslant \bar{s}_{T_2}^{(h)} \leqslant \bar{s}_{T_{12}}^{(f)},$$

d'où, tout comme dans la démonstration du théorème 1, il résulte que l'intégrabilité de f sur v entraîne celle de h sur $[c_1, c_2]$ et que leurs intégrales coïncident.

Il est évident que toute fonction f continue sur v satisfait aux hypothèses des théorèmes 1 et 1′.

2. Calcul de l'intégrale triple sur un volume compris entre deux graphes

THÉORÈME 2. *Soient ψ_1 et ψ_2 des fonctions continues sur un compact quarrable σ du plan xOy et soit $\psi_1(x, y) \leqslant \psi_2(x, y)$ pour tout* $(x, y) \in \sigma$. *Alors le corps*

$$v = \{(x, y, z) \in \mathbf{R}^3 | (x, y) \in \sigma \quad et \quad \psi_1(x, y) \leqslant z \leqslant \psi_2(x, y)\}$$

est un compact cubable. Si une fonction $f(x, y, z)$ est intégrable sur v et si pour tout point fixé $(x, y) \in \sigma$ elle est intégrable sur le segment $[\psi_1(x, y), \psi_2(x, y)]$ en tant que fonction de z, la fonction

$$g(x, y) = \int_{\psi_1(x, y)}^{\psi_2(x, y)} f(x, y, z)\, \mathrm{d}z \tag{2}$$

est intégrable sur σ et

$$\iiint_v f = \iint_\sigma g = \iint_\sigma \left(\int_{\psi_1(x, y)}^{\psi_2(x, y)} f(x, y, z)\, \mathrm{d}z \right) \mathrm{d}x\, \mathrm{d}y. \tag{3}$$

Démonstration. En tant que fonctions continues sur un compact, ψ_1 et ψ_2 sont bornées sur σ. Soient c et d des nombres quelconques satisfaisant aux inégalités $c < \psi_1(x, y)$ et $d > \psi_2(x, y)$ pour tout $(x, y) \in \sigma$. Posons

$$v^* = \sigma \times [c, d]$$

(fig. 47). Il suit du théorème 11.2.1 que v^* est un corps cubable, de sorte que sa frontière est de volume nul. Par conséquent, sa partie

$$\mathscr{F} = \{(x, y, z) \in \mathbf{R}^3 | (x, y) \in \operatorname{Fr} \sigma \quad \text{et} \quad \psi_1(x, y) \leqslant z \leqslant \psi_2(x, y)\}$$

possède la même propriété. Or la frontière de v est la réunion de $\mathscr{F}$ et des graphes des fonctions ψ_1 et ψ_2. Ces fonctions étant intégrables sur σ (voir théorème 11.3.2), leurs graphes sont des corps de volume nul (voir corollaire 2 du théorème 11.3.1). Pour cette raison, la frontière de v a un volume nul et, par suite, v est cubable. En outre, puisque les fonctions ψ_1 et ψ_2 sont continues et σ est fermé, on obtient (tout comme dans la démonstration du théorème 11.5.2) que v est fermé. Donc, v est un compact cubable. Posons

$$f^*(x, y, z) = \begin{cases} f(x, y, z) & \text{sur} \quad v, \\ 0 & \text{sur} \quad v^* \setminus v \end{cases}$$

et $v' = [v^* \setminus v]$, de sorte que $v' = v_1' \cup v_2'$ où

$$v_1' = \{(x, y, z) \in \mathbf{R}^3 | (x, y) \in \sigma \quad \text{et} \quad c \leqslant z \leqslant \psi_1(x, y)\},$$
$$v_2' = \{(x, y, z) \in \mathbf{R}^3 | (x, y) \in \sigma \quad \text{et} \quad \psi_2(x, y) \leqslant z \leqslant d\}.$$

En raisonnant comme dans la démonstration du théorème 11.5.2, on s'assure que f^* satisfait aux hypothèses du théorème 1 (à condition de remplacer f par f^* et v par v^*) et que

$$\iiint_v f = \iiint_{v^*} f^* = \iint_\sigma \left(\int_c^d f^*(x, y, z)\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y =$$

$$= \iint_\sigma \left(\int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x, y, z)\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y.$$

EXEMPLE 1. Calculons l'intégrale $\iiint_v z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ où v est l'intersection des boules $x^2 + y^2 + z^2 \leqslant 1$ et $x^2 + y^2 + (z - 1)^2 \leqslant 1$ (fig. 48). La surface de la moitié supérieure de la première boule et celle de la moitié inférieure de la deuxième boule sont respectivement exprimées par les équations

$$z = \sqrt{1 - x^2 - y^2} \quad \text{et} \quad z = 1 - \sqrt{1 - x^2 - y^2}$$

et se coupent suivant la circonférence

$$\left\{(x, y, z) \in \mathbf{R}^3 \mid x^2 + y^2 = \frac{3}{4} \quad \text{et} \quad z = \frac{1}{2}\right\}.$$

C'est pourquoi, v est une « lentille » limitée par les graphes des fonctions

$$\psi_1(x, y) = 1 - \sqrt{1 - x^2 - y^2} \quad \text{et} \quad \psi_2(x, y) = \sqrt{1 - x^2 - y^2}$$

définies sur le disque $\sigma = \left\{(x, y) \in \mathbf{R}^2 \mid x^2 + y^2 \leqslant \frac{3}{4}\right\}$. Par conséquent, on

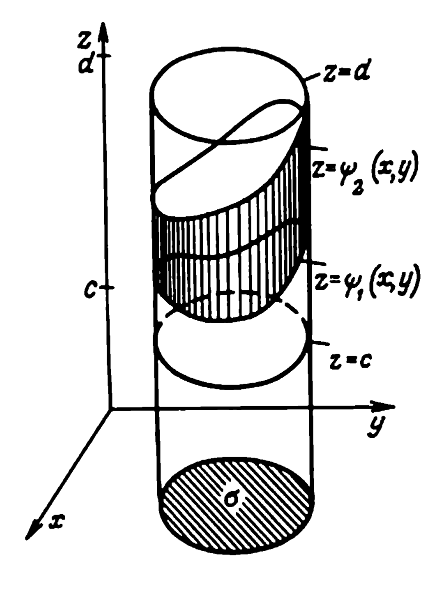


Fig. 47

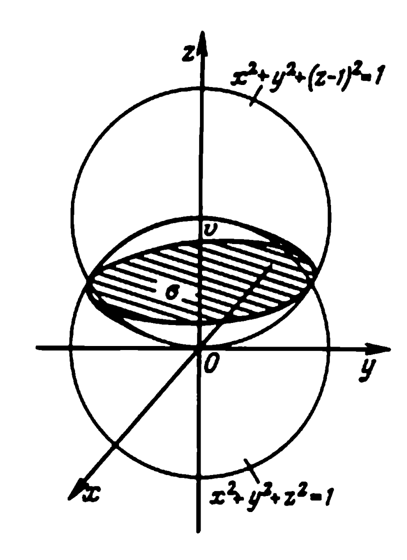


Fig. 48

a suivant le théorème 2 :

$$\iiint_v z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iint_\sigma \left(\int_{1-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} z\,\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y =$$

$$= \iint_\sigma \left(\sqrt{1-x^2-y^2} - \frac{1}{2} \right) \mathrm{d}x\,\mathrm{d}y.$$

Le passage de la dernière intégrale en coordonnées polaires (voir n° 11.4.5) donne

$$\iiint_v z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^{2\pi} \left(\int_0^{\frac{\sqrt{3}}{2}} \left(\sqrt{1-r^2} - \frac{1}{2} \right) r\,\mathrm{d}r \right) \mathrm{d}\varphi = \frac{5}{24}\,\pi.$$

Le théorème 2 ramène le calcul de l'intégrale triple au calcul de l'intégrale simple et puis de l'intégrale double. On verra plus tard que la valeur de l'intégrale triple peut, dans certaines conditions, être trouvée au moyen du calcul consécutif de trois intégrales simples (d'où la dénomination « intégrale triple »).

Désignons les points $(x, y) \in \mathbf{R}^2$ par une seule lettre u et écrivons (u, z) au lieu de (x, y, z) respectivement.

LEMME 1. *Soient ψ_1 et ψ_2 des fonctions continues sur un compact σ de $\mathbf{R}^2$ et $\psi_1(u) \leqslant \psi_2(u)$ pour tout $u \in \sigma$. Si f est une fonction continue définie sur le corps*

$$v = \{(u, z) \in \mathbf{R}^3 \mid u \in \sigma \text{ et } \psi_1(u) \leqslant z \leqslant \psi_2(u)\},$$

la fonction

$$g(u) = \int_{\psi_1(u)}^{\psi_2(u)} f(u, z)\,\mathrm{d}z$$

(définie pour tout $u \in \sigma$) est continue sur σ.

Démonstration. Tout comme dans la démonstration du théorème 2, on choisit des nombres c et d satisfaisant aux inégalités $c < \psi_1(u)$ et $d > \psi_2(u)$ pour tout $u \in \sigma$ et on pose $v^* = \sigma \times [c, d]$, de sorte que $v \subset v^*$. Par hypothèse, f est une fonction continue sur v. On prolonge f à une fonction continue F sur v^*. Cela peut être fait par plusieurs moyens, par exemple on pose, pour tout point $(u, z) \in v^*$,

$$F(u, z) = \begin{cases} f(u, \psi_1(u)) & \text{si } z < \psi_1(u), \\ f(u, z) & \text{si } \psi_1(u) \leqslant z \leqslant \psi_2(u) \quad \text{(i.e. } (u, z) \in v), \\ f(u, \psi_2(u)) & \text{si } z > \psi_2(u). \end{cases}$$

En effet, il est facile de vérifier que dans ce cas on a

$$F(u, z) = f(u, \min\{\max\{z, \psi_1(u)\}, \psi_2(u)\}),$$

de sorte que F est continue en tant que composée des fonctions continues*) ψ_1, ψ_2, max, min et f. Soient u_0 un point fixe et u un point variable du compact σ. Comme F coïncide avec f sur v, il vient

$$(4)\quad g(u) - g(u_0) = \int_{\psi_1(u)}^{\psi_2(u)} F(u, z)\,dz - \int_{\psi_1(u_0)}^{\psi_2(u_0)} F(u_0, z)\,dz =$$

$$= \int_{\psi_1(u_0)}^{\psi_2(u_0)} [F(u, z) - F(u_0, z)]\,dz + \int_{\psi_1(u)}^{\psi_1(u_0)} F(u, z)\,dz + \int_{\psi_2(u_0)}^{\psi_2(u)} F(u, z)\,dz$$

où les trois dernières intégrales ont un sens puisque les bornes d'intégration sont dans $[c, d]$ et donc le point (u, z) de chaque expression à intégrer appartient à v^* (bien qu'il puisse ne pas appartenir à v) et F (à la différence de f) est définie et continue partout sur v^*. Vu que v^* est un compact, F est uniformément continue sur v^*. Il s'ensuit que pour $\varepsilon > 0$ donné il existe un $\delta_1 > 0$ tel que $\|u - u_0\| < \delta_1$ implique

$$|F(u, z) - F(u_0, z)| < \frac{\varepsilon}{3[\psi_2(u_0) - \psi_1(u_0) + \varepsilon]}$$

pour tout $z \in [\psi_1(u_0), \psi_2(u_0)]$. En outre, F est bornée sur v^*, $|F(u, z)| < C$ pour tout $(u, z) \in v^*$. Puisque ψ_1 et ψ_2 sont continues, il existe pour le même ε un $\delta_2 > 0$ tel que $\|u - u_0\| < \delta_2$ entraîne

$$|\psi_1(u_0) - \psi_1(u)| < \frac{\varepsilon}{3C} \quad \text{et} \quad |\psi_2(u) - \psi_2(u_0)| < \frac{\varepsilon}{3C}.$$

Il résulte alors de (4) que si $\|u - u_0\| < \min\{\delta_1, \delta_2\}$, on a

$$|g(u) - g(u_0)| \leqslant \int_{\psi_1(u_0)}^{\psi_2(u_0)} |F(u, z) - F(u_0, z)|\,dz +$$

$$+ \left| \int_{\psi_1(u)}^{\psi_1(u_0)} |F(u, z)|\,dz \right| + \left| \int_{\psi_2(u_0)}^{\psi_2(u)} |F(u, z)|\,dz \right| \leqslant$$

---

*) La continuité des fonctions max et min découle des formules

$$\max\{v, w\} = \frac{v + w + |v - w|}{2}, \qquad \min\{v, w\} = \frac{v + w - |v - w|}{2}.$$

$$\leqslant \frac{\varepsilon}{3[\psi_2(u_0) - \psi_1(u_0) + \varepsilon]}[\psi_2(u_0) - \psi_1(u_0)] + C|\psi_1(u_0) - \psi_1(u)| +$$

$$+ C|\psi_2(u) - \psi_2(u_0)| < \frac{\varepsilon}{3} + C\frac{\varepsilon}{3C} + C\frac{\varepsilon}{3C} = \varepsilon,$$

de sorte que g est continue en (tout) u_0 du compact σ.

THÉORÈME 3. *Soient φ_1 et φ_2 des fonctions continues sur le segment $[a, b]$ et $\varphi_1(x) \leqslant \varphi_2(x)$ pour tout $x \in [a, b]$ et soient ψ_1 et ψ_2 des fonctions continues sur l'ensemble*

$$\sigma = \{(x, y) \in \mathbf{R}^2 | a \leqslant x \leqslant b, \quad \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$$

et $\psi_1(x, y) \leqslant \psi_2(x, y)$ pour tout $(x, y) \in \sigma$. Alors le corps

$$v = \{(x, y, z) \in \mathbf{R}^3 | (x, y) \in \sigma \quad \textit{et} \quad \psi_1(x, y) \leqslant z \leqslant \psi_2(x, y)\}$$

est un compact cubable et toute fonction continue f sur v vérifie la formule

$$\iiint_v f = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \left(\int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x, y, z)\,\mathrm{d}z \right) \mathrm{d}y \right) \mathrm{d}x. \tag{5}$$

Démonstration. Comme σ est un compact quarrable d'après le théorème 11.5.2, le théorème 2 affirme que v est un compact cubable. Etant continue sur v, la fonction f satisfait aux hypothèses du théorème 2 et vérifie pour cette raison la formule (3). En vertu du lemme 1, la fonction g définie par la formule (2) est continue, si bien qu'elle satisfait aux hypothèses du théorème 11.5.2 concernant f. Par conséquent, on a d'après ce théorème :

$$\iint_\sigma g = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} g(x, y)\,\mathrm{d}y \right) \mathrm{d}x,$$

ce qui donne, avec les formules (2) et (3), la formule (5).

EXEMPLE 2. Calculons l'intégrale $\iiint_v xy\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ où v est un tétraèdre limité par les plans de coordonnées et par le plan $x + y + z = 1$ (fig. 49). On a

$$v = \{(x, y, z) \in \mathbf{R}^3 | 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1 - x,\ 0 \leqslant z \leqslant 1 - x - y\}.$$

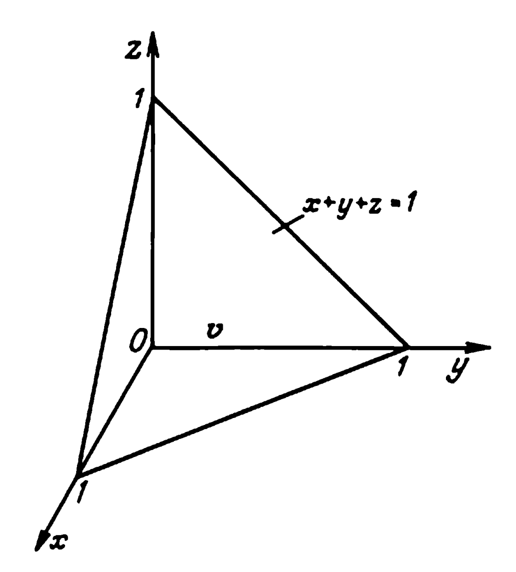


Fig. 49

Par suite, d'après le théorème 3,

$$\iiint_v xy\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1 x\left(\int_0^{1-x} y\left(\int_0^{1-x-y} \mathrm{d}z\right)\mathrm{d}y\right)\mathrm{d}x =$$

$$= \int_0^1 x\left(\int_0^{1-x} y(1-x-y)\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 x\,\frac{(1+x)^3}{6}\,\mathrm{d}x\,;$$

en faisant la substitution $u = 1 - x$, on obtient

$$\iiint_v xy\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1 (1-u)\,\frac{u^3}{6}\,\mathrm{d}u = \frac{1}{120}\,.$$

3. Expression de l'intégrale triple par les intégrales sur les sections planes

Le théorème 1′ admet la généralisation suivante (fig. 50).

THÉORÈME 4. *Soit v un compact cubable dans $\mathbf{R}^3$ ayant les propriétés suivantes :*

1) *la projection de v sur l'axe z est le segment $[c_1, c_2]$ de cet axe,*

2) *les sections de v par des plans parallèles au plan xOy, et par suite, leurs projections*

$$\sigma_z = \{(x, y) | (x, y, z) \in v\} \qquad (c_1 \leqslant z \leqslant c_2)$$

sur ce plan, sont des compacts quarrables.

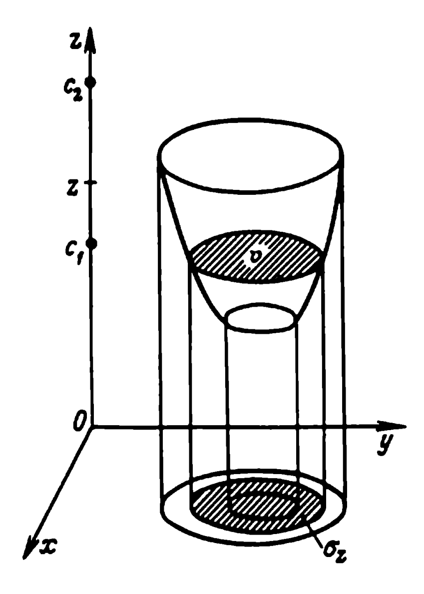


Fig. 50

Si $f(x, y, z)$ est une fonction intégrable sur v et si pour tout $z \in$ $\in [c_1, c_2]$ fixé elle est intégrable sur σ_z en tant que fonction du point (x, y), la fonction

$$h(z) = \iint_{\sigma_z} f(x, y, z)\, dx\, dy$$

est intégrable sur le segment $[c_1, c_2]$ *et*

$$\iiint_v f = \int_{c_1}^{c_2} \left(\iint_{\sigma_z} f(x, y, z)\, dx\, dy \right) dz. \tag{6}$$

Démonstration. Vu que le compact v est borné, ses projections sur les axes x et y sont respectivement contenues dans les segments $[a_1, a_2]$ et $[b_1, b_2]$ de ces axes, de sorte que v est contenu dans le parallélépipède rectangle

$$v^* = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2].$$

Posons

$$f^*(x, y, z) = \begin{cases} f(x, y, z) & \text{sur} \quad v, \\ 0 & \text{sur} \quad v^* \setminus v \end{cases}$$

et $v' = [v^* \setminus v]$; la fonction f^* est intégrable sur v puisqu'elle y coïncide avec f. D'autre part, f^* ne peut différer de zéro sur v' qu'en des points appartenant à la frontière de $v^* \setminus v$. Or

$$\text{Fr}\,(v^* \setminus v) = \text{Fr}\,(v^* \cap \text{C}_{v^*} v) \subset \text{Fr}\, v^* \cup \text{Fr}\,(\text{C}_{v^*} v) = \text{Fr}\, v^* \cup \text{Fr}\, v,$$

et de plus vol (Fr v^*) = 0 puisque Fr v^* est formée des faces du parallélépipède v^*, et vol (Fr v) = 0 puisque v est cubable. Par conséquent, Fr ($v^* \setminus v$) est de volume nul, et donc, f^* est intégrable sur v' et $\iiint_{v} f^* = 0$, de sorte que

$$\iiint_{v^*} f^* = \iiint_{v} f^* + \iiint_{v} f^* = \iiint_{v} f^* = \iiint_{v} f.$$

De façon analogue, on démontre que quel que soit $z \in [c_1, c_2]$ la fonction $f^*(x, y, z)$ est intégrable sur le rectangle

$$\sigma^* = [a_1, a_2] \times [b_1, b_2]$$

en tant que fonction du point (x, y) et que

$$\iint_{\sigma^*} f^*(x, y, z)\, \mathrm{d}x\, \mathrm{d}y = \iint_{\sigma_z} f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y = h(z).$$

Mais d'après le théorème 1′

$$\iiint_{v^*} f^* = \int_{c_1}^{c_2} \left(\iint_{\sigma^*} f^*(x, y, z)\, \mathrm{d}x\, \mathrm{d}y \right) \mathrm{d}z.$$

Par conséquent,

$$\iiint_{v} f = \int_{c_1}^{c_2} h(z)\, \mathrm{d}z.$$

EXEMPLE 3. Calculons maintenant l'intégrale $\iiint_{v} z\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z$ envisagée dans l'exemple 1 en utilisant la formule (6). La projection de la « lentille » v sur l'axe z (voir fig. 48) est le segment [0, 1], de sorte que la formule (6) donne

$$\iiint_{v} z\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = \int_0^1 z \left(\iint_{\sigma_z} \mathrm{d}x\, \mathrm{d}y \right) \mathrm{d}z = \int_0^1 z \text{ aire } \sigma_z\, \mathrm{d}z.$$

Or σ_z est un disque de rayon

$$r = \begin{cases} \sqrt{1 - (1 - z)^2} & \text{si} \quad 0 \leqslant z \leqslant 1/2, \\ \sqrt{1 - z^2} & \text{si} \quad 1/2 \leqslant z \leqslant 1. \end{cases}$$

Donc,

$$\iiint_{v} z\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = \pi \int_0^{\frac{1}{2}} z(2z - z^2)\, \mathrm{d}z + \pi \int_{\frac{1}{2}}^{1} z(1 - z^2)\, \mathrm{d}z = \frac{5}{24}\, \pi.$$

Le théorème 4 implique directement le PRINCIPE DE CAVALIERI. *Si les sections des compacts cubables v' et v'' par tout plan parallèle à un plan donné Π sont quarrables et de même aire, v' et v'' possèdent le même volume.*

En effet, il suffit de prendre Π pour le plan xOy et supposer que $f \equiv 1$ dans le théorème 4.

REMARQUE. Puisque les coordonnées x, y, z sont équivalentes, les théorèmes de ce paragraphe restent en vigueur dans toute permutation de coordonnées (les formules et démonstrations étant respectivement modifiées). Par exemple, le théorème 4 admet aussi la rédaction suivante.

Soit v un compact cubable dans $\mathbf{R}^3$ ayant les propriétés suivantes :

1) *la projection de v sur l'axe x est le segment $[a_1, a_2]$ de cet axe,*

2) *les sections de v par des plans parallèles au plan yOz et donc leurs projections*

$$\sigma_x = \{(y, z) | (x, y, z) \in v\} \qquad (a_1 \leqslant x \leqslant a_2)$$

sur ce plan sont des compacts quarrables.

Si $f(x, y, z)$ est une fonction intégrable sur v et si pour tout $x \in \in [a_1, a_2]$ fixé elle est intégrable sur σ_x en tant que fonction du point (y, z), la fonction

$$h(x) = \iint_{\sigma_x} f(x, y, z)\, \mathrm{d}y\, \mathrm{d}z$$

est intégrable sur $[a_1, a_2]$ et

$$\iiint_v f = \int_{a_1}^{a_2} \left(\iint_{\sigma_x} f(x, y, z)\, \mathrm{d}y\, \mathrm{d}z \right) \mathrm{d}x. \tag{6'}$$

4. Volume du corps de révolution

L'application de la formule (6′) donne une formule d'après laquelle on calcule le volume du corps de révolution.

THÉORÈME 5. *Etant donné une fonction positive $\varphi(x)$ définie et continue sur le segment $[a, b]$, le corps v balayé par son sous-graphe pendant une révolution autour de l'axe x* (fig. 51) *est cubable et son volume est égal à* $\pi \int_a^b [\varphi(x)]^2\, \mathrm{d}x$.

Démonstration. Vu que $v = \{(x, y, z) | a \leqslant x \leqslant b,\ y^2 + z^2 \leqslant [\varphi(x)]^2\}$ et φ est continue, v est un compact. Sa frontière est la surface $y^2 + z^2 = [\varphi(x)]^2$ (si l'on néglige deux disques de volume nul formés par les plans $x = a$ et $x = b$). Elle est constituée par les graphes des fonctions $\pm\sqrt{[\varphi(x)]^2 - y^2}$ ayant un ensemble de définition quarrable

$$\{(x, y) | a \leqslant x \leqslant b,\ -\varphi(x) \leqslant y \leqslant \varphi(x)\},$$

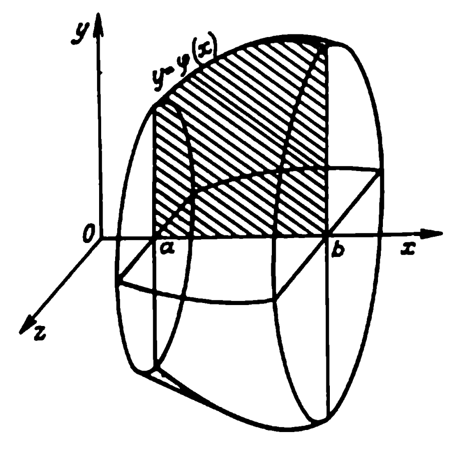


Fig. 51

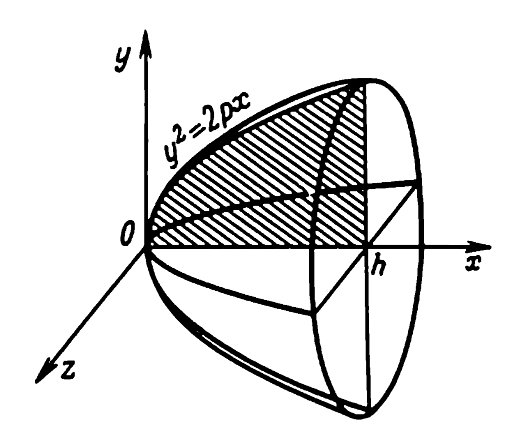


Fig. 52

et par conséquent, elle est de volume nul d'après le corollaire du théorème 11.3.1. Donc, le corps v est cubable. Sa projection sur l'axe x est le segment $[a, b]$. Par conséquent, on obtient d'après la formule (6′) :

$$\text{vol}\, v = \iiint_v \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_a^b \left(\iint_{\sigma_x} \mathrm{d}y\,\mathrm{d}z\right) \mathrm{d}x = \int_a^b \text{aire}\, \sigma_x\, \mathrm{d}x = \pi \int_a^b [\varphi(x)]^2\, \mathrm{d}x$$

puisque σ_x est le disque $\{(y, z) | y^2 + z^2 \leqslant [\varphi(x)]^2\}$ de rayon $\varphi(x)$, de sorte que aire $\sigma_x = \pi[\varphi(x)]^2$.

EXEMPLE 4. Soit v un segment du paraboloïde de révolution balayé par le segment

$$\{(x, y) | 0 \leqslant x \leqslant h,\ y^2 \leqslant 2px\}$$

de la parabole $y^2 = 2px$ tournant autour de son axe (fig. 52). D'après le théorème 5,

$$\text{vol}\, v = \pi \int_0^h 2px\, \mathrm{d}x = \pi h^2 p.$$

A noter que le volume d'un corps cylindrique ayant les mêmes base et hauteur est égal à $\pi \times 2ph \times h$. Ainsi donc, le volume d'un segment du paraboloïde de révolution est égal à la moitié du volume d'un corps cylindrique circonscrit.

§ 3. Changement de variables dans l'intégrale triple

Ce paragraphe étudie la transformation de l'intégrale triple

$$\iiint_v F(x, y, z)\, \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

par changement de variables de la forme

$$x = \varphi(u, v, w), \qquad y = \psi(u, v, w), \qquad z = \chi(u, v, w)$$

où φ, ψ, χ sont les composantes d'une application régulière ou quasi régulière de l'espace muni des coordonnées u, v, w dans l'espace muni des coordonnées x, y, z.

1. Formule générale de changement de variable

Désignons par R_{uvw} et R_{xyz} les espaces rapportés respectivement aux coordonnées cartésiennes rectangulaires u, v, w et x, y, z et identifions ces espaces à $\mathbf{R}^3$. Conformément à la définition 12.1.1, une application f de R_{uvw} dans R_{xyz} sera dite *régulière* si elle est régulière en tant qu'application de $\mathbf{R}^3$ dans $\mathbf{R}^3$, i.e. si elle est définie, continue, différentiable et injective sur un ouvert non vide D, et l'opérateur $f'(z)$ est inversible pour tout $z \in D$, i.e. $\det f'(z)$ ne s'annule nulle part sur D. Par analogie avec 12.4.2, on dit que f est *quasi régulière* sur l'adhérence $[D]$ d'un ensemble ouvert non vide D si f est définie et continûment différentiable sur un voisinage de l'ensemble $[D]$. Ceci étant, f est injective sur D, $\det f'(z) \neq 0$ pour tout $z \in D$ et, en outre, $\varkappa \cap \mathrm{Fr}\, D$ est un corps de volume nul pour tout compact cubable $\varkappa$ contenu dans $[D]$. On a les théorèmes suivants.

THÉORÈME 1. *Soit $f: R_{uvw} \to R_{xyz}$ une application définie et régulière sur un ensemble ouvert non vide D, et soit $\varkappa$ un compact cubable contenu dans $[D]$. Dans ce cas, $f(\varkappa)$ est un compact cubable, et la fonction F est intégrable sur ce compact si et seulement si la fonction $(F \circ f)\,|\det f'|$ est intégrable sur $\varkappa$. On a alors*

$$\iiint\limits_{f(\varkappa)} F = \iiint\limits_{\varkappa} (F \circ f)|\det f'|. \tag{1}$$

THÉORÈME 2. *Soit $f: R_{uvw} \to R_{xyz}$ une application définie et quasi régulière sur l'adhérence $[D]$ d'un ensemble ouvert non vide D et soit $\varkappa$ un compact cubable contenu dans $[D]$. Alors $f(\varkappa)$ est un compact cubable, et la fonction F y est intégrable si et seulement si la fonction $(F \circ f)|\det f'|$ est intégrable sur $\varkappa$. Ceci étant, on a la formule* (1).

Pour démontrer ces théorèmes, il suffit de reproduire les raisonnements faits dans la démonstration de leurs analogues bidimensionnels (voir ch. 12) en y remplaçant, bien sûr, les parallélogrammes par les parallélépipèdes, les applications affines des plans par les transformations affines de l'espace et la quarrabilité par la cubabilité.

2. Passage dans l'intégrale triple en coordonnées sphériques

La position d'un point M de l'espace de coordonnées R_{xyz} est parfaitement définie par la longueur r de son rayon vecteur $\overrightarrow{OM}$, par l'angle θ formé

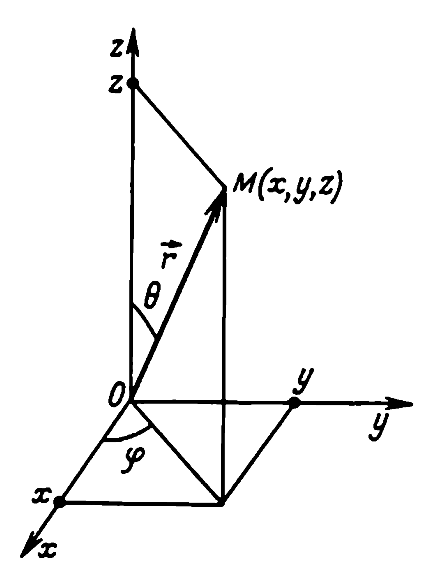


Fig. 53

par ce vecteur avec la direction positive de l'axe z, et par l'angle φ que forme la projection de $\overrightarrow{OM}$ sur le plan xOy avec la direction positive de l'axe x (fig. 53). Les nombres r, θ, φ s'appellent *coordonnées sphériques* du point M. On voit à la fig. 53 qu'elles sont liées aux coordonnées cartésiennes x, y, z de M par les égalités

$$(2) \qquad x = r \sin\theta \cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta.$$

Les seconds membres des formules (2) peuvent être envisagés en tant que composantes d'une application f de l'espace $R_{r\theta\varphi}$ (rapporté aux coordonnées *cartésiennes* r, θ, φ) dans l'espace R_{xyz} :

$$(2') \quad f_1(r, \theta, \varphi) = r\sin\theta\cos\varphi, \qquad f_2(r, \theta, \varphi) = r\sin\theta\sin\varphi,$$
$$f_3(r, \theta, \varphi) = r\cos\theta.$$

Cette application est continûment différentiable et son jacobien est

$$(3) \quad \det f'(r, \theta, \varphi) = \begin{vmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \theta} & \frac{\partial f_1}{\partial \varphi} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \theta} & \frac{\partial f_2}{\partial \varphi} \\ \frac{\partial f_3}{\partial r} & \frac{\partial f_3}{\partial \theta} & \frac{\partial f_3}{\partial \varphi} \end{vmatrix} =$$

$$= \begin{vmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} = r^2\sin\theta.$$

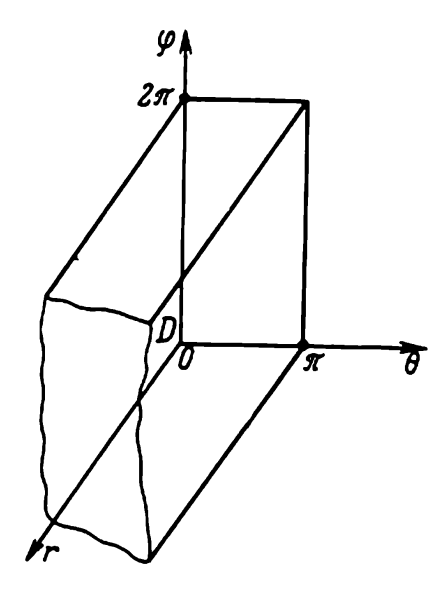


Fig. 54

Soit maintenant

$$D = \{(r, \theta, \varphi) \in R_{r\theta\varphi} | r > 0, \ 0 < \theta < \pi, \ 0 < \varphi < 2\pi\},$$

de sorte que

$$[D] = \{r, \theta, \varphi) \in R_{r\theta\varphi} | r \geqslant 0, \ 0 \leqslant \theta \leqslant \pi, \ 0 \leqslant \varphi \leqslant 2\pi\}$$

(fig. 54). Etant exprimé par les inégalités strictes, D est un ensemble ouvert. Il résulte de la formule (3) que $\det f'(r, \theta, \varphi)$ ne s'annule nulle part sur D. On voit aisément que f est injective sur D (cf. démonstration de l'injectivité de l'application étudiée dans l'exemple 12.4.1). D'autre part, f est continûment différentiable sur le voisinage $R_{r\theta\varphi}$ de l'ensemble $[D]$ et, enfin, $x \cap \operatorname{Fr} D$ est un corps de volume nul pour tout compact cubable x (ainsi que pour tout ensemble borné) dans $[D]$ puisque $\operatorname{Fr} D$ se compose de parties des plans $r = 0$, $\theta = 0$, $\theta = \pi$, $\varphi = 0$ et $\varphi = 2\pi$. Ainsi donc, f est quasi régulière sur $[D]$ et, par conséquent, il suit du théorème 2 que tout compact cubable $x \subset [D]$ et toute fonction F intégrable sur $f(x)$ vérifient la formule

$$(4) \quad \iiint\limits_{f(x)} F(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z =$$
$$= \iiint\limits_{x} F(r \sin\theta \cos\varphi, \ r \sin\theta \sin\varphi, \ r\cos\theta) \cdot r^2 \sin\theta\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\varphi.$$

Il est à noter que la définition des coordonnées sphériques implique $f([D]) = R_{xyz}$.

EXEMPLE 1. Calculons $\iiint\limits_{v} (x^2 + y^2 + z^2)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z$ où v est une boule de rayon 1 et de centre $(0, 0, 0) \in R_{xyz}$. Soit f une application de l'espace

$R_{r\theta\varphi}$ dans R_{xyz} de composantes (2′). Alors $v = f(x)$, où x est le parallélépipède $[0, 1] \times [0, \pi] \times [0, 2\pi]$ dans $R_{r\theta\varphi}$. En appliquant la formule (4) et puis la formule (5) du § 2, on obtient

$$\iiint_v (x^2 + y^2 + z^2)\,dx\,dy\,dz = \iiint_x r^4 \sin\theta\,dr\,d\theta\,d\varphi =$$

$$= \int_0^{2\pi}\left(\int_0^{\pi}\left(\int_0^1 r^4 \sin\theta\,dr\right)d\theta\right)d\varphi = \int_0^{2\pi} d\varphi \cdot \int_0^{\pi} \sin\theta\,d\theta \cdot \int_0^1 r^4\,dr = \frac{4}{5}\pi.$$

3. Passage dans l'intégrale triple en coordonnées cylindriques

On obtient les *coordonnées cylindriques* d'un point M de l'espace de coordonnées R_{xyz} à partir de ses coordonnées cartésiennes x, y, z si on remplace les coordonnées x, y de la projection de M sur le plan xOy par les coordonnées polaires r, φ (fig. 55). Ainsi donc, les coordonnées cylindriques du point M sont liées à ses coordonnées cartésiennes par les égalités :

(5) $$x = r\cos\varphi, \qquad y = r\sin\varphi, \qquad z = z.$$

Les seconds membres des formules (5) peuvent être envisagés comme composantes d'une application f de l'espace $R_{r\varphi z}$ (rapporté aux coordonnées cartésiennes r, φ, z) dans l'espace R_{xyz} :

(5′) $$f_1(r, \varphi, z) = r\cos\varphi, \qquad f_2(r, \varphi, z) = r\sin\varphi, \qquad f_3(r, \varphi, z) = z.$$

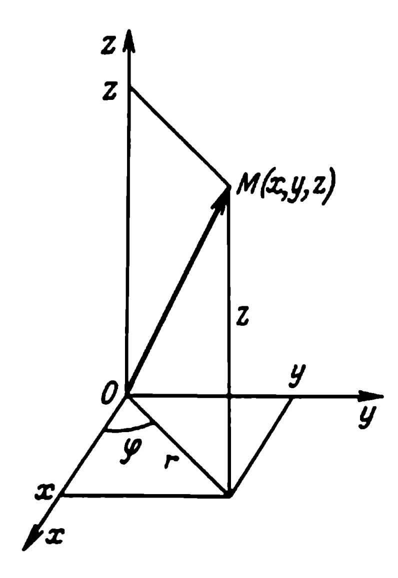


Fig. 55

Cette application est continûment différentiable et son jacobien est

$$\det f'(r, \varphi, z) = \begin{vmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r.$$

Soit maintenant

$$D = \{(r, \varphi, z) \in R_{r\varphi z} \,|\, r > 0,\ 0 < \varphi < 2\pi,\ -\infty < z < +\infty\},$$

de sorte que

$$[D] = \{(r, \varphi, z) \in R_{r\varphi z} \,|\, r \geqslant 0,\ 0 \leqslant \varphi \leqslant 2\pi,\ -\infty < z < +\infty\}$$

(fig. 56) ; D est un ensemble ouvert et $\det f'(r, \varphi, z)$ ne s'annule nulle part sur D. La démonstration de l'injectivité de l'application envisagée dans l'exemple 12.4.1 est en même temps celle de l'injectivité de f sur D. La différentiabilité continue de f sur $R_{r\varphi z}$ est déjà établie. Enfin, vu que Fr D est composée de parties des plans $r = 0$, $\varphi = 0$, $\varphi = 2\pi$ et $z = 0$, $x \cap \mathrm{Fr}\, D$ est un corps de volume nul pour tout ensemble borné x. Ainsi donc, f est quasi régulière sur $[D]$, de sorte que tout compact cubable $x \subset [D]$ et toute fonction F intégrable sur $f(x)$ vérifient la formule

$$(6) \qquad \iiint\limits_{f(x)} F(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = \iiint\limits_{x} F(r\cos\varphi, r\sin\varphi, z)\, r\, \mathrm{d}r\, \mathrm{d}\varphi\, \mathrm{d}z.$$

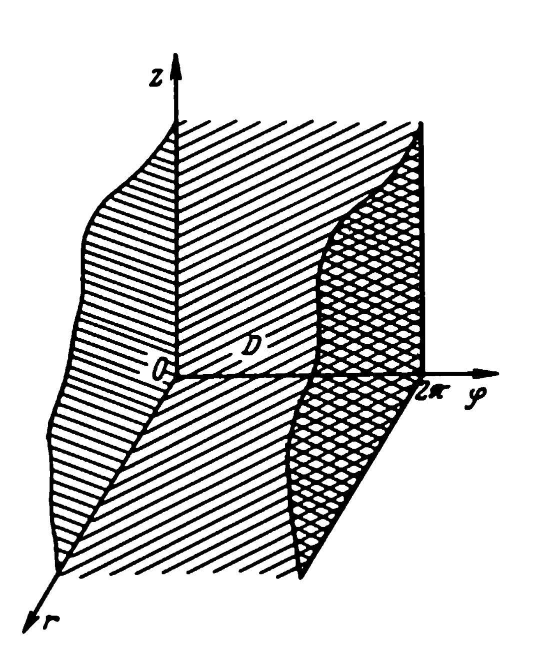


Fig. 56

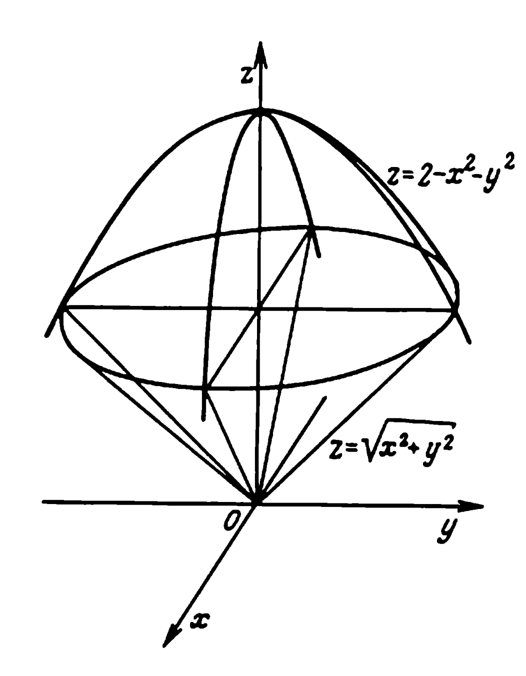


Fig. 57

EXEMPLE 2. Calculons $\iiint\limits_{v} \sqrt{x^2 + y^2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ où v est l'intersection du cône $z \geqslant \sqrt{x^2 + y^2}$ et du paraboloïde de révolution $z \leqslant 2 - x^2 - y^2$ (fig. 57). Soit f une application de l'espace $R_{r\varphi z}$ dans R_{xyz} de composantes (5′). On a alors $v = f(x)$ où

$$x = \{(r,\ \varphi,\ z) \in R_{r\varphi z} | 0 \leqslant r \leqslant 1,\ 0 \leqslant \varphi \leqslant 2\pi,\ r \leqslant z \leqslant 2 - r^2\}.$$

En appliquant la formule (6) et ensuite la formule (5) du § 2, on obtient

$$\iiint\limits_{v} \sqrt{x^2 + y^2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iiint\limits_{x} r^2\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}z =$$

$$= \int_0^{2\pi} \left(\int_0^1 \left(\int_r^{2-r^2} r^2\,\mathrm{d}z \right) \mathrm{d}r \right) \mathrm{d}\varphi = \int_0^{2\pi} \mathrm{d}\varphi \int_0^1 r^2(2 - r^2 - r)\,\mathrm{d}r = \frac{13}{30}\,\pi.$$

CHAPITRE 14

QUELQUES APPLICATIONS DES INTÉGRALES MULTIPLES

§ 1. Aire de surface

1. Exemple de Schwarz

On définit la longueur d'un arc en tant que borne supérieure des longueurs des lignes polygonales inscrites dans cet arc. Il serait naturel de définir de manière analogue l'aire d'une surface non plane comme borne supérieure des aires des surfaces polygonales inscrites. Cependant cette borne supérieure peut devenir infinie même pour des surfaces curvilignes les plus simples ! La cause en est que certaines faces du polyèdre inscrit, qui sont aussi proches que l'on veut de la surface, peuvent ne pas former avec elle de petits angles (tandis que la petite corde sous-tendant un arc d'une courbe différentiable forme de petits angles avec les tangentes à l'arc). Ainsi, le triangle dont les sommets sont situés sur une même parallèle d'un cylindre circulaire est perpendiculaire à cette surface (fig. 58). Grâce à cela on peut inscrire dans ce cylindre un polyèdre dont l'aire de surface est aussi grande que l'on veut. Voici un exemple proposé à la fin du siècle passé par le mathématicien allemand Schwarz. Considérons un cylindre circulaire droit de rayon de base 1 et de hauteur h. Divisons cette dernière en n parties égales et menons par les points de division les plans perpendiculaires à l'axe du cylindre. Ils coupent la surface cylindrique suivant les circonférences ; on obtient $n + 1$ circonférences, y compris les circonférences des bases inférieure et supérieure. Subdivisons chaque circonférence en m arcs égaux de telle sorte que les points de subdivision des circonférences voisines soient disposés en échiquier (fig. 59). Inscrivons dans le cylindre un polyèdre dont les faces sont des triangles ayant pour base la corde joignant les points voisins de la circonférence horizontale et pour sommet opposé le point de la circonférence voisine, situé au-dessus ou au-dessous du milieu de l'arc sous-tendu par la base. Ce polyèdre sera formé de $2mn$ triangles

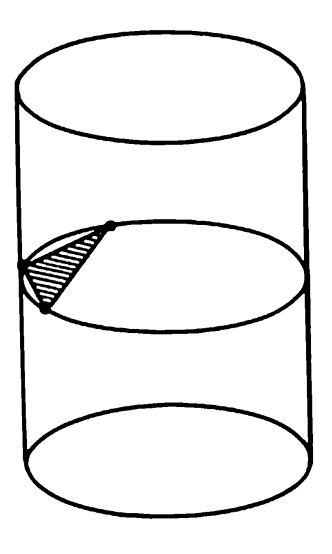


Fig. 58

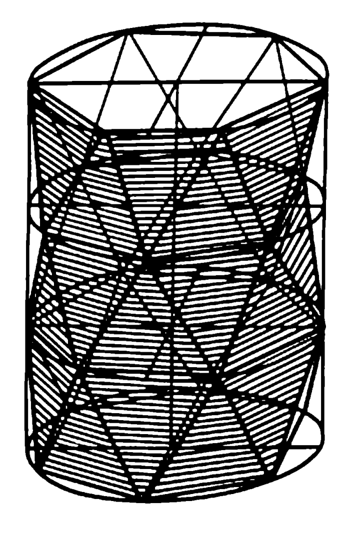


Fig. 59

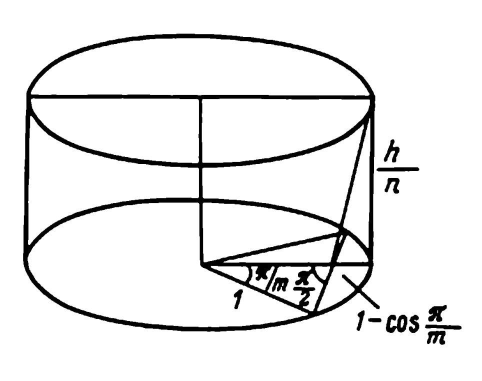


Fig. 60

congruents ayant chacun une base égale à $2\sin\frac{\pi}{m}$ et une hauteur égale à $\sqrt{\left(\frac{h}{n}\right)^2 + \left(1-\cos\frac{\pi}{m}\right)^2}$ (fig. 60), de sorte que l'aire de surface du polyèdre sera

$$2mn\sin\frac{\pi}{m}\sqrt{\left(\frac{h}{n}\right)^2 + 4\sin^2\frac{\pi}{2m}} = 2\pi\frac{\sin\frac{\pi}{m}}{\frac{\pi}{m}}\sqrt{h^2 + 4n^2\sin^2\frac{\pi}{2m}} =$$

$$= 2\pi\frac{\sin\frac{\pi}{m}}{\frac{\pi}{m}}\sqrt{h^2 + \pi^2\left(\frac{n}{m}\right)^2\left(\frac{\sin\frac{\pi}{2m}}{\frac{\pi}{2m}}\right)^2}.$$

En fixant m et en prenant n suffisamment grand, on peut rendre cette aire aussi grande que l'on veut. Elle tendra de même vers l'infini si l'on fait tendre m et n vers l'infini de telle façon que $\frac{n}{m} \to \infty$. Et si m et n tendent vers l'infini de telle sorte que $\frac{n}{m} \to c \in \mathbf{R}$, l'aire de surface tend vers $2\pi\sqrt{h^2 + \pi^2c^2}$ (i.e. vers un nombre quelconque $\geqslant 2\pi h$).

2. Longueur d'une courbe différentiable

Pour les courbes différentiables (i.e. les courbes qui admettent en chacun de leurs points une tangente dont la pente dépend continûment de ce point), la longueur peut aussi être définie comme borne *inférieure* (ou limite) des longueurs des lignes polygonales *circonscrites* (fig. 61). Cependant, ce n'est pas très rationnel pour le calcul : ayant choisi des points de tangence, on doit encore chercher les points d'intersection des tangentes voisines. Mais l'idée de base est juste : une portion suffisamment petite d'une courbe différentiable ne diffère pratiquement pas de celle d'une tangente menée en n'importe quel de ses points. Pour éviter cet inconvénient, il faut remplacer la ligne polygonale circonscrite par une « écaille » circonscrite. Considérons

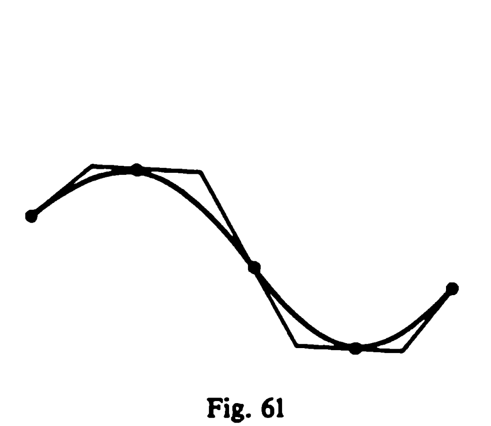


Fig. 61

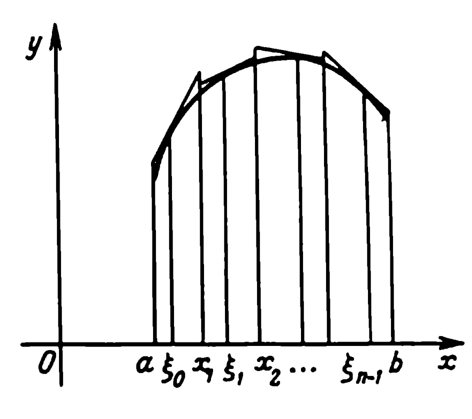


Fig. 62

à titre d'exemple le graphe d'une fonction f continûment différentiable sur le segment $[a, b]$. Subdivisons cet intervalle en segments $[x_k, x_{k+1}]$ ($k \in [0, n-1]$) et choisissons dans chacun d'eux un point ξ_k. Traçons la tangente

$$y - f(\xi_k) = f'(\xi_k)(x - \xi_k)$$

au point correspondant $(\xi_k, f(\xi_k))$ du graphe et considérons sur cette tangente un segment correspondant à $[x_k, x_{k+1}]$ (fig. 62). Soient l_k la longueur de ce segment de tangente et y_k et y_{k+1}, les ordonnées de ses extrémités. Vu que

$$y_{k+1} - y_k = f'(\xi_k)(x_{k+1} - x_k),$$

on a

$$l_k = \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2} = \sqrt{1 + f'^2(\xi_k)}(x_{k+1} - x_k).$$

Il s'ensuit que la somme des longueurs des segments de l'« écaille » construite est une somme de Riemann $\sum_{k=0}^{n-1} \sqrt{1 + f'^2(\xi_k)}\Delta x_k$ de l'intégrale

$$\int_0^b \sqrt{1 + f'^2(x)}\mathrm{d}x,$$

de sorte que cette intégrale est la limite vers laquelle tend la longueur de l' « écaille » circonscrite lorsque le pas de subdivision du segment $[a, b]$ tend vers zéro. On montre que la même intégrale est la limite des longueurs des lignes polygonales inscrites dans l'arc donné du graphe. Ainsi donc, la longueur du graphe d'une fonction de classe C^1 peut être définie en tant que limite de la longueur de l' « écaille » circonscrite.

Le calcul de l'aire d'une surface de classe C^1 sera fondé sur les mêmes principes.

3. Surfaces paramétrées

DÉFINITION 1. On appelle *surface paramétrée* toute application $f: \mathbf{R}^2 \to \mathbf{R}^3$ dont l'ensemble de définition D_f a un intérieur $]D_f[$ non vide. Les équations

$$x_1 = f_1(u_1, u_2), \quad x_2 = f_2(u_1, u_2), \quad x_3 = f_3(u_1, u_2) \quad ((u_1, u_2) \in D_f),$$

où f_1, f_2, f_3 sont les composantes de l'application f, s'appellent *équations paramétriques* de cette surface. Le graphe d'une fonction réelle f des varia-

bles réelles u, v dont l'ensemble de définition possède un intérieur non vide sera identifié à une surface donnée par les équations paramétriques

$$x = u, \quad y = v, \quad z = f(u, v) \quad ((u, v) \in D_f).$$

DÉFINITION 2. Soit f une surface paramétrée telle que : 1) l'application f est continue au point $u^0 = (u_1^0, u_2^0) \in]D_f[$ et 2) il existe un voisinage U du point u^0 contenu dans D_f tel que $f(u) \neq f(u^0)$ pour tous les points $u = (u_1, u_2) \in U$ distincts de u^0. On appelle *plan tangent* à la surface (image) f au point $M_0 = f(u^0)$ un plan Π_0 passant par M_0 et tel que la distance MN d'un point $M = f(u)$ à ce plan soit infiniment petite en comparaison avec $M_0M = \|f(u) - f(u^0)\|$ lorsque $u \to u^0$.

REMARQUE. La définition 6.3.2 du plan tangent au graphe d'une fonction réelle f des variables réelles u, v devient un cas particulier de la définition 2 si on identifie le graphe à la surface paramétrée conformément à la définition 1. Ceci étant, la condition 2) est automatiquement vérifiée puisque

$$(u, v, f(u, v)) \neq (u_0, v_0, f(u_0, v_0))$$

en général pour tous les points $(u, v) \in D_f$ distincts de (u_0, v_0).

On aura besoin de certains résultats puisés dans le cours d'algèbre et de géométrie analytique, savoir de l'*identité de Lagrange* (directement vérifiée)

$$(1) \quad (a_{11}^2 + a_{21}^2 + a_{31}^2)(a_{12}^2 + a_{22}^2 + a_{32}^2) - (a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32})^2 = \\ = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}^2 + \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}^2 + \begin{vmatrix} a_{31} & a_{32} \\ a_{11} & a_{12} \end{vmatrix}^2$$

et de deux lemmes.

LEMME 1. *Les vecteurs* e_1 *et* e_2 *de coordonnées respectives* a_{11}, a_{21}, a_{31} *et* a_{12}, a_{22}, a_{32} *sont linéairement indépendants si et seulement si la matrice*

$$(2) \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

est de rang 2, *i.e. si au moins un des déterminants*

$$(3) \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}, \quad \begin{vmatrix} a_{31} & a_{32} \\ a_{11} & a_{12} \end{vmatrix}$$

n'est pas nul.

Démonstration. La condition figurant dans le lemme est équivalente à ce que le second membre de la formule (1) soit strictement positif, i.e.

à ce que soit vérifiée l'inégalité

$$(a_{11}^2 + a_{21}^2 + a_{31}^2)(a_{12}^2 + a_{22}^2 + a_{32}^2) - (a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32})^2 > 0.$$

Mais en vertu du théorème 1.1.1 et de la remarque à ce théorème, cette inégalité est vérifiée si et seulement si les vecteurs e_1, e_2 sont linéairement indépendants.

LEMME 2. *Soient d_1 et d_2 des vecteurs dans $\mathbf{R}^2$ de coordonnées* 1, 0 *et* 0, 1 *issus du point $u^0 = (u_1^0, u_2^0)$, e_1 et e_2 des vecteurs dans $\mathbf{R}^3$ de coordonnées a_{11}, a_{21}, a_{31} et a_{12}, a_{22}, a_{32} issus du point $x^0 = (x_1^0, x_2^0, x_3^0)$ et f^0 une application de $\mathbf{R}^2$ dans $\mathbf{R}^3$ donnée par les équations paramétriques*

$$(4)\ X_i = x_i^0 + a_{i1}(u_1 - u_1^0) + a_{i2}(u_2 - u_2^0)\quad (u = (u_1, u_2) \in \mathbf{R}^2,\ i = 1, 2, 3).$$

L'image Π_0 par f^0 est un plan si (2) *est une matrice de rang* 2, *et alors f^0 est une application affine de $\mathbf{R}^2$ sur Π_0 ; Π_0 est une droite si* (2) *est une matrice de rang* 1, *i.e. si les déterminants* (3) *sont nuls mais au moins un élément de la matrice n'est pas nul ; enfin, Π_0 est un point, si* (2) *est une matrice de rang* 0, *i.e. si tous les éléments de la matrice sont nuls.*

Démonstration. Conformément au lemme 1, les vecteurs e_1, e_2 sont linéairement indépendants si la matrice (2) est de rang 2. Supposons que cette condition soit satisfaite, si bien que (e_1, e_2) est un repère (fig. 63). Vu que le repère (d_1, d_2) est obtenu du repère canonique dans $\mathbf{R}^2$ par translation de l'origine des coordonnées au point $u^0 = (u_1^0, u_2^0)$, le point $u = (u_1, u_2) \in \mathbf{R}^2$ a les coordonnées $u_1 - u_1^0$, $u_2 - u_2^0$ par rapport au repère (d_1, d_2). L'application f^0 envoie u en un point $X = (X_1, X_2, X_3) \in \mathbf{R}^3$ défini par les formules (4). Or le point X a les mêmes coordonnées $u_1 - u_1^0$, $u_2 - u_2^0$ par rapport au repère (e_1, e_2) puisque le vecteur de coordonnées $X_i - x_i^0$ issu de l'origine x^0 du repère (e_1, e_2) est une combinaison linéaire des vecteurs e_1, e_2 à coefficients $u_1 - u_1^0$, $u_2 - u_2^0$. Ainsi, f^0 applique tout point $u \in \mathbf{R}^2$

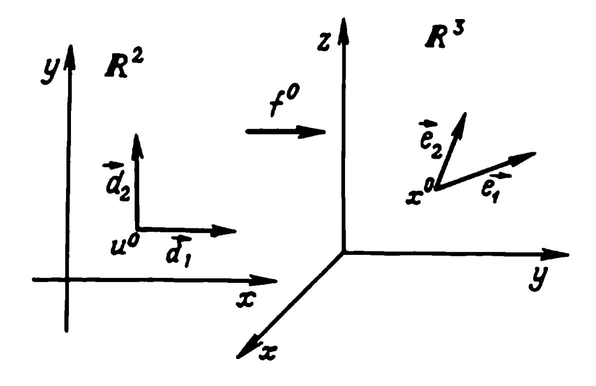


Fig. 63

en un point $X \in \Pi_0$ ayant par rapport au repère (e_1, e_2) les mêmes coordonnées que le point u a par rapport au repère (d_1, d_2). Cela signifie que Π_0 est un plan et que f^0 est une application affine de $\mathbf{R}^2$ sur Π_0. Si la matrice (2) est de rang 1, de sorte que les vecteurs e_1 et e_2 sont linéairement dépendants et au moins un de ces vecteurs n'est pas nul, Π_0 dégénère en une droite, support de ces vecteurs. Enfin, si la matrice (2) est de rang 0, i.e. $e_1 = e_2 = 0$, Π_0 dégénère en un point, soit x^0.

THÉORÈME 1. *Soit $f : \mathbf{R}^2 \to \mathbf{R}^3$ une application de composantes f_1, f_2, f_3, qui est différentiable au point $u^0 \in]D_f[$. Si sa matrice jacobienne*

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \\ f_{31} & f_{32} \end{pmatrix},$$

où $f_{ij} = \dfrac{\partial f_i}{\partial u_j}$ $(i = 1, 2, 3, j = 1, 2)$, possède en ce point le rang 2, i.e. au moins un des déterminants

$$\begin{vmatrix} f_{11}(u^0) & f_{12}(u^0) \\ f_{21}(u^0) & f_{22}(u^0) \end{vmatrix}, \quad \begin{vmatrix} f_{21}(u^0) & f_{22}(u^0) \\ f_{31}(u^0) & f_{32}(u^0) \end{vmatrix}, \quad \begin{vmatrix} f_{31}(u^0) & f_{32}(u^0) \\ f_{11}(u^0) & f_{12}(u^0) \end{vmatrix}$$

n'est pas nul, l'image par f a au point $M_0 = f(u^0)$ un plan tangent. Ce plan est l'image Π_0 par une application de $\mathbf{R}^2$ dans $\mathbf{R}^3$ donnée par les équations paramétriques

$$X_i = f_i(u^0) = f_{i1}(u^0)(u_1 - u_1^0) + f_{i2}(u^0)(u_2 - u_2^0) \quad (u = (u_1, u_2) \in \mathbf{R}^2, \ i = 1, 2, 3). \tag{5}$$

Démonstration. L'application f satisfait à la condition 1) de la définition 2 en vertu de sa différentiabilité au point u^0.

Démontrons que la condition 2) est également satisfaite. Suivant le théorème 7.2.5, l'application f ainsi que ses composantes f_i sont différentiables au point u^0, de sorte qu'en vertu du corollaire au théorème 6.3.5

$$x_i - x_i^0 = g_{i1}\Delta u_1^0 + g_{i2}\Delta u_2^0 \qquad (i = 1, 2, 3), \tag{6}$$

où $x_i = f_i(u)$, $x_i^0 = f_i(u^0)$, $\Delta u_j^0 = u_j - u_j^0$ $\quad (j = 1, 2)$ et

$$g_{ij} \to f_{ij}(u^0) \text{ lorsque } u \to u^0. \tag{7}$$

C'est pourquoi,

$$\|f(u) - f(u^0)\|^2 = \sum_{i=1}^{3} (x_i - x_i^0)^2 = A(\Delta u_1^0)^2 + 2B\Delta u_1^0\Delta u_2^0 + C(\Delta u_2^0)^2 \tag{8}$$

où $A = \sum_{i=1}^{3} g_{i1}^2$, $B = \sum_{i=1}^{3} g_{i1}g_{i2}$, $C = \sum_{i=1}^{3} g_{i2}^2$. Posons

$$E = \sum_{i=1}^{3} f_{i1}^2(u^0), \quad F = \sum_{i=1}^{3} f_{i1}(u^0) f_{i2}(u^0), \quad G = \sum_{i=1}^{3} f_{i2}^2(u^0).$$

En vertu de l'hypothèse du théorème, on a

$$\begin{vmatrix} f_{11}(u^0) & f_{12}(u^0) \\ f_{21}(u^0) & f_{22}(u^0) \end{vmatrix}^2 + \begin{vmatrix} f_{21}(u^0) & f_{22}(u^0) \\ f_{31}(u^0) & f_{32}(u^0) \end{vmatrix}^2 + \begin{vmatrix} f_{31}(u^0) & f_{32}(u^0) \\ f_{11}(u^0) & f_{12}(u^0) \end{vmatrix}^2 > 0,$$

ou, ce qui est la même chose, $EG - F^2 > 0$ d'après l'identité de Lagrange (1). Mais en vertu de (7) on a $AC - B^2 \to EG - F^2$ lorsque $u \to u^0$. Par conséquent, il existe un voisinage U du point u^0 tel que $AC - B^2 > 0$ et, par suite, $A > 0$ et $C > 0$ pour tout $u \in U$. Il découle alors de la formule (8) que pour tout $u \in U$

$$\|f(u) - f(u^0)\|^2 = \frac{1}{A} [(A\Delta u_1^0 + B\Delta u_2^0)^2 + (AC - B^2)(\Delta u_2^0)^2] \geqslant$$
$$\geqslant \frac{1}{A} (AC - B^2)(\Delta u_2^0)^2,$$
$$\|f(u) - f(u^0)\|^2 = \frac{1}{C} [(B\Delta u_1^0 + C\Delta u_2^0)^2 + (AC - B^2)(\Delta u_1^0)^2] \geqslant$$
$$\geqslant \frac{1}{C} (AC - B^2)(\Delta u_1^0)^2$$

et, par conséquent,

(9) $$\|f(u) - f(u^0)\|^2 \geqslant \varrho[(\Delta u_2^0)^2 + (\Delta u_1^0)^2] = \varrho \|u - u^0\|^2$$

où $\varrho = \frac{1}{2}\left(\frac{AC - B^2}{A} + \frac{AC - B^2}{C}\right) > 0$. Mais il résulte de (9) que $f(u) \neq f(u^0)$ pour tout $u \in U$ distinct de u^0, i.e. la condition 2) de la définition 2 est aussi vérifiée. Il reste à montrer que les points $X = (X_1, X_2, X_3)$ satisfaisant aux équations paramétriques (5) forment un plan tangent à l'image par f au point u^0. En vertu du lemmc 2 et de l'hypothèse du théorème, ils forment effectivement un plan ; désignons-le par Π_0. Soit MN la distance du point $M = f(u)$ à ce plan. Vu que $X \in \Pi_0$, on a $MN \leqslant MX$, d'où en vertu de (5), (6) et de l'inégalité de Cauchy,

$$MN^2 \leqslant MX^2 = \sum_{i=1}^{3} (X_i - x_i)^2 =$$
$$= \sum_{i=1}^{3} \{[f_{i1}(u_0) - g_{i1}]\Delta u_1^0 + [f_{i2}(u^0) - g_{i2}]\Delta u_2^0\}^2 \leqslant$$
$$\leqslant \sum_{i=1}^{3} \{[f_{i1}(u) - g_{i1}]^2 + [f_{i2}(u^0) - g_{i2}]^2\} \|u - u^0\|^2.$$

En prenant en considération (9) pour tout $u \in U$ distinct de u^0, on a donc

$$\left(\frac{MN}{M_0M}\right)^2 = \left(\frac{MN}{\|f(u) - f(u_0)\|}\right)^2 \leqslant \frac{MX^2}{\|f(u) - f(u^0)\|^2} \leqslant$$
$$\leqslant \frac{1}{\varrho} \sum_{i=1}^{3} \{[f_{i1}(u^0) - g_{i1}]^2 + [f_{i2}(u^0) - g_{i2}]^2\}.$$

Or, en vertu de (7), la dernière somme tend vers zéro lorsque $u \to u^0$, et

$$\varrho \to \frac{1}{2}\left(\frac{EG - F^2}{E} + \frac{EG - F^2}{G}\right) \neq 0.$$

Par conséquent, $\dfrac{MN}{M_0M} \to 0$, c.q.f.d.

4. Aire d'une surface différentiable

DÉFINITION 3. On appelle *surface différentiable* tout ensemble $S \subset \mathbf{R}^3$ de la forme $S = f(\sigma)$ où f est une application continûment différentiable de $\mathbf{R}^2$ dans $\mathbf{R}^3$ et σ, un compact quarrable non vide contenu dans D_f et qui coïncide avec l'adhérence $[]\sigma[]$ de son intérieur $]\sigma[$. Ceci étant, f est injective sur $]\sigma[$ (de sorte que $f(]\sigma[)$ peut être identifiée à $f|_{]\sigma[}$) et la matrice jacobienne de l'application f est de rang 2 partout dans $]\sigma[$ (et, par conséquent, S a un plan tangent partout dans $f(]\sigma[)$).

Soient f_1, f_2, f_3 les composantes de l'application f et soient $f_{ij} = \dfrac{\partial f_i}{\partial u_j}$ $(i = 1, 2, 3, j = 1, 2)$. Posons

$$(10) \qquad E(u) = \sum_{i=1}^{3} f_{i1}^2(u), \qquad F(u) = \sum_{i=1}^{3} f_{i1}(u) f_{i2}(u),$$
$$G(u) = \sum_{i=1}^{3} f_{i2}^2(u) \qquad (u \in D_j).$$

Comme $[]\sigma][= \sigma \neq \varnothing$, on a $]\sigma[\neq \varnothing$. Soit T une partition de σ en éléments quarrables $\sigma_1, \ldots, \sigma_m$ ayant des points communs avec $]\sigma[$. Cette partition ayant un diamètre aussi petit que l'on veut, peut être obtenue de la façon suivante. Vu que $]\sigma[\neq \varnothing$, il existe des carrés de rang n (voir n° 11.1.1) ayant une intersection non vide avec $]\sigma[$ et, puisque σ est borné, leur nombre est fini. Soient $Q_1, \ldots, Q_m$ ces carrés et soit $\sigma_k = Q_k \cap \sigma$ $(k \in [1, m])$. En tant qu'intersections des compacts quarrables, σ_k sont des compacts quarrables, et $]\sigma_k[\cap]\sigma_l[= \varnothing$ pour $k \neq l$ car $]Q_k[\cap]Q_l[= \varnothing$. En outre, $\sigma' = \bigcup_{k=1}^{m} \sigma_k = \sigma$; en effet, il est évident que $\sigma' \subset \sigma$ et, d'autre part,

$\sigma = [] \sigma [] \subset \sigma'$ car $]\sigma[\subset \left(\bigcup_{k=1}^{m} Q_k\right) \cap \sigma = \sigma'$ et σ' est fermé. Ainsi donc, σ_k forment une partition de σ et $\sigma_k \cap]\sigma[= Q_k \cap]\sigma[\neq \varnothing$.

Choisissons dans chaque σ_k de la partition T un point $u^k \in]\sigma[$ et soit Π_k l'image d'une application $f_k : \mathbf{R}^2 \to \mathbf{R}^3$ définie par les équations paramétriques

$$X_i = f_i(u^k) + f_{i1}(u^k)(u_1 - u_1^k) + f_{i2}(u^k)(u_2 - u_2^k)$$
$$(u = (u_1, u_2) \in \mathbf{R}^2, \ i = 1, 2, 3).$$

Comme $u^k \in]\sigma[$, la matrice de l'application f est de rang 2 au point u^k, de sorte que, suivant le théorème 1, Π_k est un plan tangent à S au point $M_k = f(u^k)$ et, en vertu du lemme 2, f^k est une application affine de $\mathbf{R}^2$ sur Π_k. Il s'ensuit d'après le théorème 12.2.2 que l'image $f^k(\Phi)$ de chaque figure quarrable $\Phi \subset \mathbf{R}^2$ est quarrable et le rapport (aire $f^k(\Phi)$/aire Φ) est constant. Soient d_1 et d_2 des vecteurs dans $\mathbf{R}^2$ de coordonnées 1, 0 et 0, 1 issus du point u^k et soient e_1 et e_2 des vecteurs dans $\mathbf{R}^3$ de coordonnées $f_{i1}(u^k)$ et $f_{i2}(u^k)$ ($i = 1, 2, 3$) issus du point M_k, de sorte que

$$|e_1|^2 = E(u^k), \quad \langle e_1, e_2 \rangle = F(u^k), \quad |e_2|^2 = G(u^k).$$

Conformément à la démonstration du lemme 2, f^k applique chaque point $u \in \mathbf{R}^2$ en un point $X \in \Pi_k$ ayant par rapport au repère (e_1, e_2) les mêmes coordonnées que u possède par rapport au repère (d_1, d_2). Il en résulte en particulier que le carré construit sur le repère (d_1, d_2) et ayant l'aire 1 a pour image un parallélogramme P construit sur le repère (e_1, e_2). Par conséquent, aire $f^k(\Phi)$/aire Φ = aire $P/1$ = aire P, de sorte qu'en particulier

$$\text{aire } f^k(\sigma_k) = \text{aire } P \cdot \text{aire } \sigma_k.$$

Or

$$\text{aire } P = |e_1||e_2| \sin (\widehat{e_1, e_2}) = |e_1||e_2|\sqrt{1 - \cos^2 (\widehat{e_1, e_2})} =$$
$$= |e_1||e_2| \sqrt{1 - \left(\frac{\langle e_1, e_2 \rangle}{|e_1||e_2|}\right)^2} = \sqrt{|e_1|^2|e_2|^2 - \langle e_1, e_2 \rangle^2} =$$
$$= \sqrt{E(u^k)G(u^k) - F^2(u^k)}.$$

Par conséquent,

$$\text{aire } f^k(\sigma_k) = \sqrt{E(u^k)G(u^k) - F^2(u^k)} \cdot \text{aire } \sigma_k \qquad (k \in [1, n]).$$

Donc, l'aire totale des plaquettes $f^k(\sigma_k)$ qui forment une « écaille» recouvrant la surface S (fig. 64) est la somme de Riemann

$$\sum_{k=1}^{n} \sqrt{E(u^k)G(u^k) - F^2(u^k)} \Delta\sigma_k$$

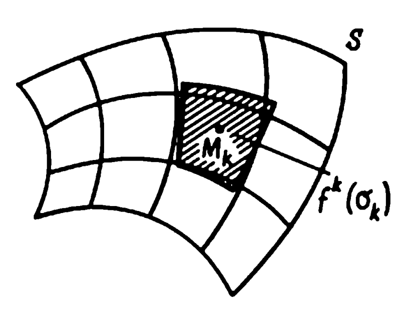


Fig. 64

de l'intégrale $\iint\limits_{\sigma} \sqrt{E(u)G(u) - F^2(u)}\,\mathrm{d}u_1\,\mathrm{d}u_2$ et, par suite, elle tend vers cette intégrale si le diamètre de T tend vers zéro. La dernière intégrale est prise par définition pour l'aire de la surface $f(\sigma)$.

EXEMPLE 1. Calculons l'aire $\mathscr{P}$ d'une partie du paraboloïde hyperbolique $z = xy$, découpée par la surface cylindrique $x^2 + y^2 = a^2$. On a une application f du compact $\sigma : x^2 + y^2 \leqslant a^2$ de $\mathbf{R}^2$ dans $\mathbf{R}^3$ de composantes $f_1(x, y) = x$, $f_2(x, y) = y$, $f_3(x, y) = xy$. Ainsi donc, on obtient, en vertu de (10) et en posant $u = (x, y)$,

$$E(u) = \sum_{i=1}^{3} f_{i1}^2(u) = 1 + y^2, \quad G(u) = \sum_{i=1}^{3} f_{i2}(u) = 1 + x^2,$$

$$F(u) = \sum_{i=1}^{3} f_{i1}(u) f_{i2}(u) = xy.$$

L'aire recherchée est égale par définition à l'intégrale $\iint\limits_{\sigma} \sqrt{1 + x^2 + y^2}\,\mathrm{d}x\,\mathrm{d}y$.

L'application de composantes $x = r\cos\varphi$, $y = r\sin\varphi$ du plan de coordonnées (r, φ) dans le plan de coordonnées (x, y) est quasi régulière sur l'adhérence $[D]$ de l'ensemble ouvert non vide $D = \{(r, \varphi) | r > 0, 0 < \varphi < 2\pi\}$. Par conséquent,

$$\iint\limits_{\sigma} \sqrt{1 + x^2 + y^2}\,\mathrm{d}x\,\mathrm{d}y = \iint\limits_{\sigma} \sqrt{1 + r^2}\, r\,\mathrm{d}r\,\mathrm{d}\varphi,$$

où $\sigma' = \{(r, \varphi) | 0 \leqslant r \leqslant a,\ 0 \leqslant \varphi \leqslant 2\pi\}$. Donc,

$$\mathscr{P} = \iint\limits_{\sigma} \sqrt{1 + r^2}\, r\,\mathrm{d}r\,\mathrm{d}\varphi =$$

$$= \int_0^{2\pi} \left(\frac{1}{2} \int_0^a (1 + r^2)^{1/2} \mathrm{d}(1 + r^2) \right) \mathrm{d}\varphi = \frac{2}{3}\,\pi((1 + a^2)^{3/2} - 1).$$

§ 2. Centres de masse

On appelle *centre de masse* d'un système fini de masses ponctuelles m_k concentrées aux points $P_k(x_k, y_k)$ $(k = 1, \ldots, n)$ du plan un point de coordonnées

$$x_c = \frac{\sum m_i x_i}{\sum m_i}, \qquad y_c = \frac{\sum m_i y_i}{\sum m_i}.$$

Admettons que la masse est continûment répartie sur une figure plane. Pour trouver son centre il faut subdiviser la figure en un nombre fini de parties, remplacer chaque partie par l'un quelconque de ses points où l'on concentre toute la masse de cette partie et trouver le centre du système obtenu des masses ponctuelles. Il est évident que plus la dimension des parties de subdivision est petite, plus les résultats de calcul sont précis ; on obtient la valeur exacte lorsque le diamètre de subdivision a pour limite zéro.

Calculons le centre de masse d'une figure plane S ayant une densité superficielle continue $\varrho(x, y)$. Considérons une partition T de S en éléments quarrables σ_k d'aire $\Delta\sigma_k$ et choisissons dans chaque σ_k un point (x_k, y_k). Admettons que la densité sur σ_k entier est égale à la densité ϱ_k au point choisi et que la masse de σ_k est concentrée en ce point. Le centre du système fini des masses ponctuelles a alors les coordonnées suivantes

$$x_T = \frac{\sum x_k \varrho_k \Delta\sigma_k}{\sum \varrho_k \Delta\sigma_k}, \qquad y_T = \frac{\sum y_k \varrho_k \Delta\sigma_k}{\sum \varrho_k \Delta\sigma_k}.$$

La valeur exacte des coordonnées du centre de masse d'une figure plane s'obtient lorsque le diamètre de T a pour limite zéro, i.e.

$$x_c = \lim_{d_T \to 0} x_T = \frac{\iint\limits_S x\varrho \,\mathrm{d}\sigma}{\iint\limits_S \varrho \,\mathrm{d}\sigma},$$

$$y_c = \lim_{d_T \to 0} y_T = \frac{\iint\limits_S y\varrho \,\mathrm{d}\sigma}{\iint\limits_S \varrho \,\mathrm{d}\sigma}.$$

Dans le cas homogène où la densité est constante, on a

$$x_c = \frac{1}{\text{aire } S} \iint\limits_S x \,\mathrm{d}\sigma, \qquad y_c = \frac{1}{\text{aire } S} \iint\limits_S y \,\mathrm{d}\sigma.$$

Soit, en particulier, S un « trapèze curviligne » :

$$S = \{(x, y) | 0 \leqslant y \leqslant f(x), \ x \in [a, b]\}, \qquad f(x) \geqslant 0, \ x \in [a, b].$$

Alors

$$\text{aire } S = \int_a^b f(x) \,\mathrm{d}x,$$

$$x_c = \frac{1}{\int_a^b f(x)\,dx} \iint_S x\,d\sigma = \frac{1}{\int_a^b f(x)\,dx} \int_a^b x\left(\int_0^{f(x)} dy\right)dx = \frac{\int_a^b xf(x)\,dx}{\int_a^b f(x)\,dx},$$

$$y_c = \frac{1}{\int_a^b f(x)\,dx} \iint_S y\,d\sigma = \frac{1}{\int_a^b f(x)\,dx} \int_a^b \left(\int_0^{f(x)} y\,dy\right)dx = \frac{\frac{1}{2}\int_a^b f^2(x)\,dx}{\int_a^b f(x)\,dx}.$$

EXEMPLE 1. Centre de masse du demi-disque $x^2 + y^2 \leqslant a^2$, $y \geqslant 0$. Ici

$$y_c = \frac{1}{\pi a^2}\int_{-a}^{a} (a^2 - x^2)\,dx = \frac{\left(a^2 x - \frac{x^3}{3}\right)}{\pi a^2}\Bigg|_{-a}^{a} = \frac{\frac{4}{3}a^3}{\pi a^2} = \frac{4}{3\pi}a.$$

De façon analogue, les coordonnées du centre de masse d'un corps v de densité $\varrho(x, y, z)$ sont calculées à l'aide d'une intégrale triple :

$$x_c = \frac{\iiint_v x\varrho}{\iiint_v \varrho}, \quad y_c = \frac{\iiint_v y\varrho}{\iiint_v \varrho}, \quad z_c = \frac{\iiint_v z\varrho}{\iiint_v \varrho}.$$

Pour les corps homogènes (ϱ = const) on a

$$x_c = \frac{1}{\text{vol } v}\iiint_v x, \quad y_c = \frac{1}{\text{vol } v}\iiint_v y, \quad z_c = \frac{1}{\text{vol } v}\iiint_v z.$$

CHAPITRE 15

INTÉGRALE CURVILIGNE

§ 1. Chemin

1. Notion de chemin

Rappelons qu'on appelle *arc paramétré* dans $\mathbf{R}^2$ toute application continue de $\mathbf{R}$ dans $\mathbf{R}^2$ ayant pour ensemble de définition un segment non dégénéré. Soient f un arc paramétré dans $\mathbf{R}^2$, $I = D_f$ et $g = f \circ j$ où j est une application strictement croissante du segment J dans I. Puisque j est monotone et I est partout dense, j est continue ; pour cette raison, g est aussi un arc paramétré. On dira qu'il est *équivalent* à l'arc paramétré f et on le notera $g \sim f$. Cette relation entre les arcs paramétrés est une équivalence, i.e. : 1) $f \sim f$, 2) $g \sim f \Rightarrow f \sim g$ et 3) $(h \sim g$ et $g \sim f) \Rightarrow h \sim f$. En effet,

1) vu que $f = f \circ i$ où i est une application identique (et par suite, strictement croissante) du segment $D_f = I$ dans lui-même, on a $f \sim f$;

2) soit $g \sim f$, i.e. $g = f \circ j$ où j est une application strictement croissante du segment $J = D_g$ dans I ; il s'ensuit que j est inversible et j^{-1} est une application strictement croissante de I dans J ; comme $f = g \circ f^{-1}$, on a $f \sim g$;

3) soit encore $h \sim g$, de sorte qu'il existe une application strictement croissante k du segment $K = D_h$ dans J telle que $h = g \circ k$; puisque $g = f \circ j$, on obtient $h = f \circ (j \circ k)$, et comme $j \circ k$ est une application strictement croissante de K dans I, on a $h \sim f$.

Soient f et g des arcs paramétrés équivalents dans $\mathbf{R}^2$, de sorte que $g = f \circ j$ où j est une application strictement croissante du segment $J = D_g$ dans le segment $I = D_f$. Si τ parcourt J par valeurs croissantes, $t = j(\tau)$ croît en parcourant I, de sorte que $g(\tau) = f(j(\tau))$ parcourt les mêmes points et dans le même ordre que $f(t)$. En se fondant sur ce fait, on dira que les arcs f et g définissent *un même chemin* et qu'ils sont ses *représentations paramétriques*. On dira aussi que les points des arcs f et g sont les points de ce chemin. En tant qu'image de l'arc paramétré, l'ensemble des points du chemin est un compact.

Ainsi, soient f et g des arcs paramétrés dans $\mathbf{R}^2$ donnés respectivement par les équations paramétriques

$$(1) \qquad x = \cos t, \qquad y = \sin t \qquad (0 \leqslant t \leqslant 2\pi)$$

et

(2) $$x = \cos 2\pi\tau, \qquad y = \sin 2\pi\tau \qquad (0 \leqslant \tau \leqslant 1).$$

Comme $g = f \circ j$ où j est une application strictement croissante $\tau \mapsto t = 2\pi\tau$ de $[0, 1]$ dans $[0, 2\pi]$, f et g sont des représentations paramétriques du même chemin dont l'ensemble des points est une circonférence exprimée par l'équation

(3) $$x^2 + y^2 = 1.$$

Quand le paramètre τ ou t parcourt le segment de ses valeurs, le point correspondant du chemin parcourt cette circonférence à partir du même point $(1, 0)$ et dans le même sens (antihoraire) ; la seule différence consiste en ce que dans le cas (2) cela se produit plus rapidement que dans le cas (1).

LEMME 1. *Tout chemin peut être paramétré par un arc dont le paramètre parcourt un segment non dégénéré* $[c, d]$ *donné à l'avance.*

Démonstration. Soit f une représentation paramétrique du chemin L et $D_f = [a, b]$. Posons pour tout $\tau \in [c, d]$

$$j(\tau) = a + \frac{\tau - c}{d - c}(b - a).$$

Alors j est une application strictement croissante de $[c, d]$ dans $[a, b]$ et $g = f \circ j$ est une représentation paramétrique du même chemin L, qui est définie déjà sur le segment $[c, d]$.

C'est le procédé le plus simple qui porte le nom de *paramétrage linéaire* (ou plus précisément affine). Si $[c, d] = [a, b]$, on a $j(\tau) = \tau$, i.e. j est une application identique du segment $[a, b]$ sur lui-même. Il existe une infinité d'autres applications strictement croissantes de $[a, b]$ sur lui-même. En voici un exemple (qui nous sera utile plus tard):

$$j_{a,b}(t) = a + \frac{b - a}{2}\left(1 - \cos\frac{\pi(t - a)}{b - a}\right) \qquad (a \leqslant t \leqslant b).$$

En effet, $j_{a,b}(a) = a$, $j_{a,b}(b) = b$, $j_{a,b}$ est différentiable et

$$j'_{a,b}(t) = \frac{\pi}{2}\sin\frac{\pi(t - a)}{b - a}\begin{cases} = 0 & \text{si } t = a, \\ > 0 & \text{si } t \in]a, b[, \\ = 0 & \text{si } t = b, \end{cases}$$

de sorte que $j_{a,b}$ croît strictement sur $[a, b]$ et applique $[a, b]$ sur lui-même (fig. 65).

Soient maintenant f et g deux représentations paramétriques du chemin L, avec $D_f = [a, b]$ et $D_g = [c, d]$, de sorte que $g = f \circ j$ où j est une application strictement croissante de $[c, d]$ dans $[a, b]$. Comme dans ce cas $a = j(c)$, $b = j(d)$, on a $f(a) = g(c)$, $f(b) = g(d)$, si bien que ces points ne

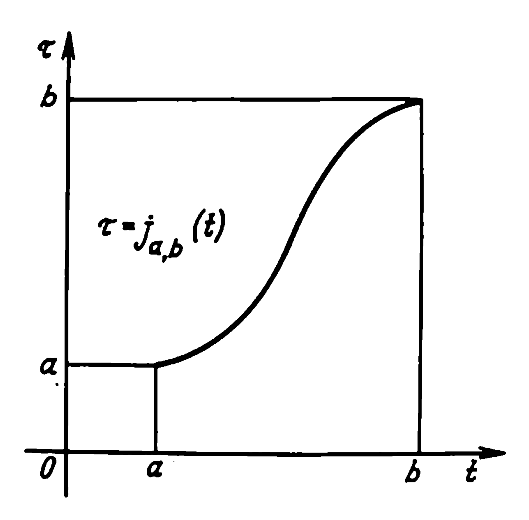


Fig. 65

dépendent pas du choix de la représentation paramétrique. Le premier point s'appelle *origine* du chemin L, le second, son *extrémité*. Si ces points coïncident, le chemin L est appelé *chemin fermé* (ou *lacet*). Ainsi, le chemin défini par les représentations paramétriques (1) et (2) est fermé : ses extrémités initiale et terminale sont présentées par le point (1, 0). D'autre part, le chemin défini par les équations paramétriques

$$(4) \qquad x = (1 - t)x_1 + tx_2 \qquad y = (1 - t)y_1 + ty_2 \qquad (0 \leqslant t \leqslant 1)$$

et dont l'ensemble des points est un segment joignant les points (x_1, y_1) et (x_2, y_2), son origine et son extrémité, n'est pas fermé si $(x_1, y_1) \neq (x_2, y_2)$.

2. Chemin inverse

Notons $\sigma_{u,v}$ la symétrie du segment $[u, v]$ par rapport à son milieu $(u + v)/2$. Elle est donnée par la formule

$$\sigma_{u,v}(t) = u + v - t \qquad (u \leqslant t \leqslant v).$$

En particulier, $\sigma_{u,v}\left(\frac{u+v}{2}\right) = \frac{u+v}{2}$, $\sigma_{u,v}(u) = v$, $\sigma_{u,v}(v) = u$. La symétrie $\sigma_{u,v}$ est une application strictement décroissante du segment $[u, v]$ sur lui-même et $\sigma_{u,v}^2 = I_{[u,v]}$, de sorte que $\sigma_{u,v}^{-1} = \sigma_{u,v}$.

Soient f et g des représentations paramétriques du chemin L, avec $D_f = [a, b]$ et $D_g = [c, d]$, de sorte que $g = f \circ j$ où j est une application strictement croissante de $[c, d]$ dans $[a, b]$. Alors

$$g \circ \sigma_{c,d} = f \circ j \circ \sigma_{c,d} = (f \circ \sigma_{a,b}) \circ (\sigma_{a,b} \circ j \circ \sigma_{c,d}).$$

Or $\sigma_{a,b} \circ j \circ \sigma_{c,d}$ est une application strictement croissante de $[c, d]$ dans $[a, b]$. Ainsi donc, $g \circ \sigma_{c,d} \sim f \circ \sigma_{a,b}$, i.e. $f \circ \sigma_{a,b}$ et $g \circ \sigma_{c,d}$ sont des représen-

tations paramétriques du même chemin. Appelons-le *chemin inverse* à L et notons-le $-L$. Il a pour origine l'extrémité $f(b) = g(d)$ du chemin L et pour extrémité l'origine $f(a) = g(c)$. Vu que $\sigma_{a,b}^2 = I_{[a,b]}$, on a $-(-L) = L$.

Ainsi, si L est un chemin défini par les équations paramétriques (1), $-L$ peut être donné par les équations paramétriques

$$x = \cos(2\pi - t) = \cos t, \quad y = \sin(2\pi - t) = -\sin t \quad (0 \leqslant t \leqslant 2\pi).$$

Le chemin inverse à celui qui est donné par les équations (4) peut encore être défini par les équations

$$x = (1-(1-t))x_1 + (1-t)x_2 = tx_1 + (1-t)x_2, \qquad y = ty_1 + (1-t)y_2 \quad (0 \leqslant t \leqslant 1).$$

3. Chemin simple

On sait que l'arc paramétré peut avoir des points doubles ou même des points multiples. Il en est de même, bien sûr, pour les chemins. Un chemin non fermé est dit *simple* s'il admet une représentation paramétrique injective. Il est évident que toutes les représentations paramétriques de ce chemin sont alors injectives. Ainsi, le chemin défini par les équations paramétriques (4) du segment joignant deux points distincts (x_1, y_1) et (x_2, y_2) est simple.

THÉORÈME 1. *L'ensemble des points d'un chemin non fermé simple est homéomorphe à un segment non dégénéré. Tout ensemble de points du plan, qui est homéomorphe à un segment non dégénéré, est un ensemble des points de deux chemins non fermés (réciproquement inverses) et de ces chemins seuls.*

Démonstration. L'ensemble E des points d'un chemin non fermé simple est l'image d'un compact (plus précisément, d'un segment non dégénéré) par une application continue injective. Or, en vertu du théorème 3.2.11, l'application inverse est aussi continue. Ainsi, E est homéomorphe à un segment non dégénéré. Soit maintenant f un homéomorphisme du segment non dégénéré $[a, b]$ sur un ensemble $F \subset \mathbf{R}^2$ et soit L le chemin qu'il définit ; f étant une application injective, L est un chemin non fermé simple. Mais les points a et b diffèrent des autres points du segment $[a, b]$ par le fait que leur absence ne viole pas la connexité (tandis qu'un segment pointé se décompose en deux parties connexes). Puisque la connexité est invariante par l'homéomorphisme, les points $f(a)$ et $f(b)$ de F possèdent la même propriété. C'est pourquoi, tout chemin non fermé simple ayant F pour son ensemble de points, doit avoir une origine $f(a)$ et une extrémité $f(b)$ ou *vice versa*. Soit g une représentation paramétrique d'un tel chemin L' ; conformément au lemme 1, on peut admettre que $D_g = [a, b]$. Si l'origine $g(a)$ du chemin L' coïncide avec $f(a)$, $j = g^{-1} \circ f$ est un homéomorphisme du segment $[a, b]$ sur lui-même qui laisse invariant le point a ; or tout homéomorphisme du segment sur lui-même est strictement monotone ; puisque $j(a) = a$, on en conclut que j est une application strictement croissante du segment $[a, b]$ sur lui-même et, comme $f = g \circ j$, on obtient $f \sim g$, de sorte que $L' = L$. Si, par contre, $g(a) = f(b)$, $j = g^{-1} \circ (f \circ \sigma_{a,b})$ est un homéomorphisme du segment $[a, b]$ sur lui-même qui laisse fixe a, i.e. j est une application strictement croissante de $[a, b]$ sur lui-même ; comme $f \circ \sigma_{a,b} = g \circ j$, on a $L' = -L$.

En particulier, tout segment orienté dans $\mathbf{R}^2$ est un ensemble des points d'un seul chemin non fermé simple ayant la même origine. En se fondant sur ce fait, on peut identifier ce segment et ce chemin. Pour la même raison, on identifie le graphe d'une fonction $g(x)$ continue sur un segment non dégénéré $[a, b]$, qui est orienté de gauche à droite, avec le chemin L défini par les équations paramétriques

$$x = t, \qquad y = g(t) \qquad (a \leqslant t \leqslant b).$$

Un chemin fermé est dit *simple* s'il admet une représentation paramétrique h définie sur un segment non dégénéré $[a, b]$ et satisfaisant à la condition suivante : si $h(x) = h(y)$ et $x < y$, on a $x = a$, $y = b$. Il est évident que toutes les représentations paramétriques de ce chemin vérifient cette condition.

THÉORÈME 2. *L'ensemble des points d'un chemin fermé simple est homéomorphe à une circonférence. Tout ensemble de points du plan, qui est homéomorphe à une circonférence, est un ensemble des points de deux chemins fermés simples (réciproquement inverses), et de ces chemins seuls, ayant une origine donnée (qui peut être représentée par tout point de l'ensemble).*

Démonstration. Soit h une représentation paramétrique d'un chemin fermé simple L ; en vertu du lemme 1, on peut poser $D_h = [0, 2\pi]$. Soit f une application du segment $[0, 2\pi]$ dans $\mathbf{R}^2$ donnée par les équations paramétriques (1) et donc, ayant pour image une circonférence C définie par l'équation (3). La restriction f_0 de f à l'intervalle $]0, 2\pi[$ est injective et, pour cette raison, inversible ; soit $g = f_0^{-1}$; g est une application de l'ensemble C_0 des points de C distincts du point $(1, 0)$ dans l'intervalle $]0, 2\pi[$.

Montrons que g est continue. En effet, soient $(x_0, y_0) \in C_0$ et $t_0 = g(x_0, y_0)$. Considérons un ε-voisinage quelconque du point t_0, contenu dans l'intervalle $]0, 2\pi[$. L'ensemble F des points de C qui correspondent aux valeurs t n'appartenant pas à ce voisinage est un compact en tant qu'image du compact $[0, t_0 - \varepsilon] \cup [t_0 + \varepsilon, 2\pi]$ par une application continue f et, par conséquent, F est un fermé. C'est pourquoi, $C \setminus F$ est un ensemble ouvert dans C. Or $(x_0, y_0) \in C \setminus F$ et $g(C \setminus F) \subset]t_0 - \varepsilon, t_0 + \varepsilon[$. Par conséquent, l'image réciproque de tout voisinage du point $g(x_0, y_0)$ par g est un voisinage du point (x_0, y_0) dans C_0, i.e. g est continue. Il en résulte en particulier que f_0 est un homéomorphisme de $]0, 2\pi[$ sur C_0.

Pour tout point $(x, y) \in C$, posons maintenant

$$\varphi(x, y) = \begin{cases} h(g(x, y)) & \text{si} \quad (x, y) \in C_0, \\ h(0) & \text{si} \quad (x, y) = (1, 0). \end{cases} \tag{5}$$

Il est évident que l'image de φ coïncide avec celle de h. Montrons que φ est continue. Pour les points $(x, y) \in C_0$ cela résulte de la continuité de g et h. D'autre part, de la façon analogue à la démonstration de la continuité de g, on peut montrer, en considérant les demi-voisinages droits du point 0 ou, respectivement, les demi-voisinages gauches du point 2π, que $g(x, y) \to 0$ lorsque (x, y) tend vers le point $(1, 0)$ en parcourant la moitié supérieure de la circonférence C et que $g(x, y) \to 2\pi$ lorsque (x, y) tend vers $(1, 0)$ en parcourant la moitié inférieure. Pour cette raison, $\varphi(x, y)$ tend vers $h(0)$ dans le premier cas et vers $h(2\pi)$ dans le deuxième. Comme le chemin L est fermé, $h(0) = h(2\pi)$. Par conséquent, $\varphi(x, y) \to \varphi(1, 0)$ lorsque $(x, y) \to (1, 0)$, i.e. φ est aussi continue au point $(1, 0)$. Enfin, φ est injective. En effet, puisque le chemin

L est simple, h est injective sur $g(C_0) =]0, 2\pi[$ et g est injective sur C_0 en tant qu'application inverse, c'est pourquoi, φ est injective sur C_0 et, en outre, $\varphi(1, 0) = h(0) \notin h(]0, 2\pi[) = \varphi(C_0)$. Ainsi, φ est une application injective continue de la circonférence C dans l'ensemble K des points du chemin L et, par conséquent, φ est un homéomorphisme de C sur K. Ceci étant, φ envoie le point $(1, 0)$ de C dans l'origine $h(0)$ du chemin L.

Soit maintenant K un ensemble quelconque de points du plan, qui est homéomorphe à une circonférence, et soit $q_0 \in K$. Etant donné que, pour deux circonférences quelconques et deux points choisis sur chacune d'elles, il existe une application homéomorphe d'une circonférence sur l'autre envoyant un point choisi dans l'autre, on trouve une application homéomorphe ψ de la circonférence C sur K envoyant le point $(1, 0)$ dans q_0. Posons $h' = \psi \circ f$. Si $h'(x) = h'(y)$, on a $f(x) = f(y)$ puisque ψ est injective, de sorte que $x = 0, y = 2\pi$ si $x < y$; d'autre part, $f(0) = f(2\pi) = (1, 0)$ implique $h'(0) = h'(2\pi) = q_0$. Ainsi donc, h' est une représentation paramétrique du chemin fermé simple L' dont l'ensemble des points est K et l'origine est q_0. Soit L un autre chemin fermé simple d'origine q_0 et de support K et soit h, sa représentation paramétrique ; en vertu du lemme 1, on peut poser que $D_h = [0, 2\pi]$ et que l'application φ définie par (5) est un homéomorphisme de C sur K d'après ce qui a été démontré plus haut. On a de plus $h = \varphi \circ f$. En effet, $\varphi \circ f_0 = h \circ g \circ f_0 = h_0$ où h_0 est la restriction de h à $]0, 2\pi[$, et $(\varphi \circ f)(0) = \varphi(1, 0) = h(0)$ et $(\varphi \circ f)(2\pi) = \varphi(1, 0) = h(2\pi)$. Soit h_0' la restriction de h' à $]0, 2\pi[$, i.e. $h_0' = \psi \circ f_0$. Puisque f_0 est un homéomorphisme de $]0, 2\pi[$ sur C_0 et que φ et ψ sont des homéomorphismes de C sur K, h_0 et h_0' sont des homéomorphismes de $]0, 2\pi[$ sur $K \setminus \{q_0\}$. Il en résulte que $j_0 = h_0^{-1} h_0'$ est un homéomorphisme et, par conséquent, une application strictement croissante ou strictement décroissante de l'intervalle $]0, 2\pi[$ sur lui-même. Ceci étant, on a $h_0' = h_0 \circ j_0$. Si j_0 est strictement croissante, posons

$$j(t) = \begin{cases} j_0(t) & \text{si} \quad t \in]0, 2\pi[, \\ 0 & \text{si} \quad t = 0, \\ 2\pi & \text{si} \quad t = 2\pi, \end{cases}$$

si bien que j est une application strictement croissante du segment $[0, 2\pi]$ sur lui-même. Comme $h'(0) = h(0)$, $h'(2\pi) = h(2\pi)$, on a $h' = h \circ j$, de sorte que $h' \sim h$, i.e. $L' = L$. Si j_0 est strictement décroissante, posons

$$j(t) = \begin{cases} \sigma_{0,2\pi} \circ j_0(t) & \text{si} \quad t \in]0, 2\pi[, \\ 0 & \text{si} \quad t = 0, \\ 2\pi & \text{si} \quad t = 2\pi, \end{cases}$$

de sorte que j est de nouveau une application strictement croissante de $[0, 2\pi]$ sur lui-même. Comme $h_0' = h_0 \circ \sigma_{0,2\pi} \circ (\sigma_{0,2\pi} \circ j_0)$ et $h'(0) = h(2\pi) = h(0) = h'(2\pi)$, on obtient $h' = (h \circ \sigma_{0,2\pi}) \circ j$; par conséquent, on a dans ce cas $h' \sim h \circ \sigma_{0,2\pi}$, i.e. $L' = -L$.

En particulier, toute circonférence de $\mathbf{R}^2$, sur laquelle on a choisi un point et un sens de parcours, est l'ensemble des points de l'unique chemin simple dont l'origine est le point choisi. En se fondant sur ce fait, on identifiera ces circonférences et ce chemin.

4. Composition des chemins

Supposons que l'extrémité du chemin L serve de l'origine du chemin M. On peut alors composer de L et de M un nouveau chemin. En effet, soient f et $\bar{g}$ des représentations paramétriques respectives des chemins L et M, avec $D_f = [a, b]$ et $D_{\bar{\gamma}} = [\bar{a}, \bar{b}]$, de sorte que $f(b) = \bar{g}(\bar{a})$. Prenons

un nombre $c > b$ et posons $g = \bar{g} \circ j$ où j est une application affine croissante de $[b, c]$ dans $[\bar{a}, \bar{b}]$:

$$j(t) = \bar{a} + \frac{t - b}{c - b}(\bar{b} - \bar{a}). \tag{6}$$

Soit g une autre représentation paramétrique du chemin M. On a maintenant $f(b) = g(b)$, et les fonctions f et g peuvent être « collées » de façon élémentaire à condition de poser

$$h(t) = \begin{cases} f(t) & \text{si} \quad a \leqslant t \leqslant b, \\ g(t) & \text{si} \quad b \leqslant t \leqslant c. \end{cases} \tag{7}$$

Au point $t = b$ les deux définitions sont compatibles, et h est continue (car elle est continue à gauche en tant que f et à droite en tant que g) ; h est une représentation paramétrique d'un chemin N. Il est aisé de voir que le même chemin peut être obtenu pour n'importe quel choix de représentations paramétriques des chemins L et M. On dira que le chemin N est le *composé des chemins* L et M et on le notera $L + M$.

La notion de composé des chemins se généralise par récurrence à toute suite finie $(L_i)_{i \in [1,n]}$ de chemins tels que l'extrémité du chemin L_i coïncide avec l'origine du chemin L_{i+1} pour tout entier $i \in [1, n - 1]$; si le composé de $n - 1$ chemins est déjà défini, on entend par composé des chemins $L_1, \ldots, L_{n-1}, L_n$ (satisfaisant à la condition mentionnée) le chemin

$$(L_1 + \ldots + L_{n-1}) + L_n. \tag{8}$$

Si les chemins L_i sont des segments, leur composé s'appelle *ligne polygonale*.

5. Chemin de classe C^1

On appelle *chemin de classe* C^1 le chemin admettant une représentation paramétrique continûment différentiable. Le segment (4) en sert d'exemple élémentaire. Notons que la différentiabilité peut disparaître lors du passage à une autre représentation paramétrique. Mais il est très important que cette représentation continûment différentiable existe.

THÉORÈME 3. *Le composé des chemins de classe C^1 est un chemin de classe C^1.*

Démonstration. Soient L et M deux chemins de classe C^1, de sorte qu'ils possèdent des représentations paramétriques continûment différentiables f et $\bar{g}$, avec $D_f = [a, b]$, $D_{\bar{g}} = [\bar{a}, \bar{b}]$ et $f(b) = \bar{g}(\bar{a})$. Si $c > b$ et j est une application affine (6) de $[b, c]$ dans $[\bar{a}, \bar{b}]$, $g = \bar{g} \circ j$ est aussi continûment différentiable. Mais h, définie par les conditions (7), peut ne pas être différentiable au point b (h possède en b les dérivées à gauche et à droite, qui peuvent ne pas coïncider). Collons alors les arcs f et g d'une

façon plus régulière. Posons

$$\bar{h}(t) = \begin{cases} f(j_{a,b}(t)) & \text{si} \quad a \leqslant t \leqslant b, \\ g(j_{b,c}(t)) & \text{si} \quad b \leqslant t \leqslant c, \end{cases}$$

où

$$j_{u,v}(t) = u + \frac{v-u}{2}\left(1 - \cos\frac{\pi(t-u)}{v-u}\right) \qquad (u \leqslant t \leqslant v)$$

(voir n° 1) ; $f \circ j_{a,b}$ et $g \circ j_{b,c}$ sont aussi des représentations paramétriques des chemins L et M. L' application $\bar{h}$ est continûment différentiable sur les intervalles $[a, b[$ et $]b, c]$. Elle est aussi continûment différentiable au point b. En effet, $\bar{h}$ a pour composantes

$$\bar{h}_i(t) = \begin{cases} f_i(j_{a,b}(t)) & \text{si} \quad a \leqslant t \leqslant b, \\ g_i(j_{b,c}(t)) & \text{si} \quad b \leqslant t \leqslant c, \end{cases}$$

où f_i sont les composantes de l'application f et g_i celles de l'application g. Puisque f_i et g_i sont différentiables au point b, $\bar{h}_i$ possède en b les dérivées à gauche et à droite. Comme $j'_{a,b}(b) = j'_{b,c}(b) = 0$, on a

$$\bar{h}'_{i\,\text{gauche}}(b) = f'_i(b) j'_{a,b}(b) = 0, \qquad \bar{h}'_{i\,\text{droite}}(b) = g'_i(b) j'_{b,c}(b) = 0.$$

Par conséquent, les dérivées $\bar{h}'_i(b)$ existent et sont nulles, d'où $\bar{h}'(b)$ existe et est nulle. En outre, puisque

$$\bar{h}'_i(t) = f'_i(j_{a,b}(t)) j_{a,b}(t) \to f'_i(b) \cdot 0 = 0 = \bar{h}'_i(b)$$

lorsque $t \to b$ par valeurs inférieures et, de façon analogue, $\bar{h}'_i(t) \to \bar{h}'_i(b)$ lorsque $t \to b$ par valeurs supérieures, la dérivée $\bar{h}'$ est aussi continue au point b. Donc, le théorème est démontré pour le composé de deux chemins. Il en résulte par récurrence que cette assertion est valable pour le composé de n'importe quel nombre de chemins de classe C^1.

COROLLAIRE. *Toute ligne polygonale est un chemin de classe C^1.*

En particulier, la frontière du triangle, du carré, etc., est un chemin de classe C^1 si l'on choisit d'une façon appropriée les origines de ses composantes rectilignes, i.e. elle admet une représentation paramétrique continûment différentiable (bien que certains de ses points soient « anguleux »).

Il faut noter en conclusion que *si L est un chemin de classe C^1, il en est de même du chemin $-L$.* Ceci résulte de la définition du chemin inverse (voir n° 2) et de la différentiabilité continue de la transformation symétrique du segment.

6. Longueur du chemin

THÉORÈME 4. *Toutes les représentations paramétriques d'un même chemin possèdent la même longueur.*

Démonstration. Soient f et g des représentations paramétriques du chemin L, de sorte que $g = f \circ j$ où j est une application strictement croissante du segment $D_g = [c, d]$ dans le segment $D_f = [a, b]$. L'application j engendre une bijection de l'ensemble $\mathscr{T}_{[c,d]}$ de toutes les subdivisions du segment $[c, d]$ sur $\mathscr{T}_{[a,b]}$. A savoir, si T' est une subdivision du segment $[c, d]$:

$$c = t'_0 < t'_1 < \dots < t'_{n-1} < t'_n = d,$$

les points $t_i = j(t'_i)$ vont dans l'ordre croissant de $t_0 = a$ jusqu'à $t_n = b$ et engendrent ainsi une subdivision T du segment $[a, b]$; notons-la $j(T')$; on a alors $T' = j^{-1}(T)$. Chaque subdivision $T \in \mathscr{T}_{[c,d]}$ engendre une ligne polygonale inscrite dans l'arc g ; la longueur de cette ligne est

$$l_{T'} = \sum_{k=0}^{n-1} \sqrt{[g_1(t'_{k+1}) - g_1(t'_k)]^2 + [g_2(t'_{k+1}) - g_2(t'_k)]^2},$$

où g_1, g_2 sont les composantes de l'application g. De façon analogue, $T = j(T')$ engendre une ligne polygonale inscrite dans l'arc f et ayant la longueur

$$l_T = \sum_{k=0}^{n-1} \sqrt{[f_1(t_{k+1}) - f_1(t_k)]^2 + [f_2(t_{k+1}) - f_2(t_k)]^2},$$

où f_1, f_2 sont les composantes de l'application f. Or $g = f \circ j$ implique

$$f_i(t_k) = (f_i \circ j)(t'_k) = g_i(t'_k) \qquad (i = 1, 2 ; k = 0, 1, \dots, n-1),$$

d'où $l_{T'} = l_T$. Par conséquent,

$$\sup_{T' \in \mathscr{T}_{[c,d]}} l_{T'} = \sup_{T \in \mathscr{T}_{[a,b]}} l_T,$$

i.e. les arcs f et g sont d'une même longueur.

Il est naturel de dire que la longueur commune aux représentations paramétriques du chemin L est la *longueur du chemin* L et que L est un *chemin rectifiable* si sa longueur est finie.

§ 2. Notion d'intégrale curviligne et ses propriétés fondamentales

1. Formes différentielles sur le plan

On appelle *forme différentielle* sur un domaine $D \subset \mathbf{R}^2$ toute famille de fonctions linéaires homogènes sur $\mathbf{R}^2$, dépendant du paramètre (x, y) parcourant D.

Toute fonction linéaire homogène sur $\mathbf{R}^2$ est de la forme

$$Ph + Qk,$$

où P et Q sont des constantes et (h, k) parcourt $\mathbf{R}^2$. C'est pourquoi, la forme différentielle sur D est définie par un couple de fonctions $P(x, y)$, $Q(x, y)$ sur D et a pour expression

(1) $$l(x, y)(h, k) = P(x, y)h + Q(x, y)k.$$

Donc, si F est une fonction différentiable sur D, sa différentielle $\mathrm{d}F(x, y)$ est une forme différentielle sur D donnée par la formule

$$\mathrm{d}F(x, y)(h, k) = F'_x(x, y)h + F'_y(x, y)k.$$

En particulier, pour les fonctions de coordonnées $F(x, y) = x$ et $F(x, y) = y$, on a respectivement

$$\mathrm{d}x(h, k) = h, \qquad \mathrm{d}y(h, k) = k.$$

En portant ces égalités dans (1), on obtient

$$l(x, y)(h, k) = P(x, y)\mathrm{d}x(h, k) + Q(x, y)\mathrm{d}y(h, k)$$

pour tout $(h, k) \in \mathbf{R}^2$, i.e.

$$l(x, y) = P(x, y)\mathrm{d}x + Q(x, y)\mathrm{d}y. \tag{2}$$

Telle est l'expression générale de la forme différentielle sur un domaine plan. En particulier

$$\mathrm{d}F(x, y) = F'_x(x, y)\mathrm{d}x + F'_y(x, y)\mathrm{d}y. \tag{3}$$

REMARQUE. On démontre en théorie des intégrales définies que toute forme différentielle $f(x)\mathrm{d}x$ possède une primitive sur l'intervalle où f est continue, i.e. qu'elle s'écrit $F'(x)\mathrm{d}x$. L'assertion analogue pour les formes différentielles sur les domaines plans n'est pas vraie en général. Plus loin, on établira les conditions suffisantes que doivent vérifier P et Q pour que la forme (2) soit une différentielle exprimée par (3).

2. Notion d'intégrale curviligne

Le problème type de physique ramenant à une intégrale curviligne est le calcul du travail produit par un champ de forces stationnaire pour le déplacement d'un point matériel. Nous allons considérer le cas d'un champ plan.

Le champ stationnaire plan est un domaine plan D (domaine de définition du champ) en tout point (x, y) duquel est défini un vecteur $\vec{F}(x, y)$ (le plus souvent une force) qui ne dépend pas du temps (pour cette raison, le champ est dit stationnaire). Soient $P(x, y)$ et $Q(x, y)$ ses composantes. Il faut évaluer le travail R effectué par ce champ de forces pour le déplacement d'un point matériel de masse unitaire le long d'une trajectoire contenue dans D.

Cas 1. La trajectoire est un segment de droite orienté $\vec{s}$, et $\vec{F}(x, y)$ est un vecteur constant colinéaire au vecteur $\vec{s}$. On pose, par définition, $R = Fs$ où $s = |\vec{s}|$ (longueur du chemin) et $F = |\vec{F}|$ si $\vec{F}$ et $\vec{s}$ sont de même sens, et $F = -|\vec{F}|$ s'ils sont de sens opposés.

Cas 2. Il se distingue du cas 1 par le seul fait que $\vec{F}$ et $\vec{s}$ ne sont pas obligatoirement colinéaires. On pose alors, par définition,

$$R = \langle \vec{F}, \vec{s} \rangle = |\vec{F}|\cdot|\vec{s}|\cdot\cos(\vec{s}, \vec{F}),$$

i.e. $R = P\Delta x + Q\Delta y$ où Δx et Δy sont les coordonnées du vecteur $\vec{s}$, et P et Q celles du vecteur $\vec{F}$. (Le travail n'est produit que par la projection orthogonale de la force $\vec{F}$ sur la trajectoire $\vec{s}$.)

Cas 3 (général). On décompose la trajectoire L en éléments $L_1, \ldots, L_n$ et on choisit un point (ξ_k, η_k) arbitraire sur chaque L_k. Le travail produit le long de L_k est approximativement égal à

$$P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k,$$

où Δx_k et Δy_k sont les différences des coordonnées de l'extrémité et de l'origine du chemin L_k. Dans ce cas, le travail produit le long de toute la trajectoire L est approximativement égal à

$$\sum_k P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k.$$

Par définition, la valeur exacte R du travail produit le long de toute la trajectoire L est égale à la limite de la dernière somme (si cette limite existe) lorsque les éléments de subdivision de L deviennent infiniment petits.

Cela permet d'introduire la nouvelle espèce de l'intégrale sous une forme mathématique pure. Soient donnés :

1) une forme différentielle $P(x, y)\mathrm{d}x + Q(x, y)\mathrm{d}y$ sur un domaine plan D ; on supposera partout plus loin que P et Q sont continus ;

2) un chemin $L \subset D$, avec une représentation paramétrique f ; f_1, f_2 sont ses composantes et $[a, b] = D_f$.

On envisage toutes les subdivisions possibles (T, Ξ) du segment $[a, b]$ par des points

$$a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b \qquad \text{et} \qquad \tau_k \in [t_k, t_{k+1}]$$
$$(k = 0, 1, \ldots, n).$$

Pour chaque couple (T, Ξ) on écrit la *somme de Riemann*

$$R_{T,\Xi}^{(f)} = \sum_{k=0}^{n-1} P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k,$$

où

$$\xi_k = f_1(\tau_k), \qquad \eta_k = f_2(\tau_k),$$
$$\Delta x_k = f_1(t_{k+1}) - f_1(t_k), \qquad \Delta y_k = f_2(t_{k+1}) - f_2(t_k).$$

Si $R_{T,\Xi}^{(f)}$ tend vers le nombre R lorsque $d_T \to 0$, on dit que R est l'*intégrale de la forme différentielle* $P\,\mathrm{d}x + Q\,\mathrm{d}y$ *le long du chemin* L et on la note

$$\int_L P(x, y)\mathrm{d}x + Q(x, y)\mathrm{d}y \qquad \left(\text{ou } \int_L P\,\mathrm{d}x + Q\,\mathrm{d}y\right).$$

La forme elle-même est dite *intégrable*.

Montrons que *l'intégrale curviligne* $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y$ *ne dépend pas du choix de la représentation paramétrique f du chemin L.* En effet, soit g une autre représentation paramétrique de ce chemin et soit $D_g = [c, d]$, de sorte que $g = f \circ j$ où j est une application strictement croissante de $[c, d]$ dans $[a, b]$. L'application j engendre une correspondance biunivoque entre les subdivisions (T', Ξ') et (T, Ξ) des segments $[c, d]$ et $[a, b]$; plus précisément, si t'_k, τ'_k sont des points de la subdivision (T', Ξ') du segment $[c, d]$, les points $t_k = j(t'_k)$ et $\tau_k = j(\tau'_k)$ forment une subdivision (T, Ξ) du segment $[a, b]$ et *vice versa*. Les couples (T, Ξ) et (T', Ξ') engendrent les sommes de Riemann $R^{(f)}_{T,\Xi}$ et $R^{(g)}_{T',\Xi'}$ identiques. En effet, notons F une application de D dans $\mathbf{R}^2$ de composantes P, Q. En gardant les notations utilisées pour la définition de l'intégrale curviligne, on a alors

$$P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k = P(f(\tau_k))[(f_1(t_{k+1}) - f_1(t_k)] + \\ + Q(f(\tau_k))[f_2(t_{k+1}) - f_2(t_k)] = \langle F(f(\tau_k)), f(t_{k+1}) - f(t_k)\rangle$$

et c'est pourquoi

$$R^{(f)}_{T,\Xi} = \sum_{k=0}^{n-1} \langle F(f(\tau_k)), f(t_{k+1}) - f(t_k)\rangle = \\ = \sum_{k=0}^{n-1} \langle F(f(j(\tau'_k))), f(j(t'_{k+1})) - f(j(t'_k))\rangle = \\ = \sum_{k=0}^{n-1} \langle F(g(\tau'_k)), g(t'_{k+1}) - g(t'_k)\rangle = R^{(g)}_{T,\Xi}\,.$$

Soit maintenant $R^{(f)}_{T,\Xi} \to R$ lorsque $d_T \to 0$, i.e.

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall(T, \Xi))(d_T < \delta \Rightarrow |R^{(f)}_{T,\Xi} - R| < \varepsilon).$$

Vu que j est uniformément continue sur $[c, d]$, il existe un $\eta > 0$ tel que

$$|t'_{k+1} - t'_k| < \eta \Rightarrow |j(t'_{k+1}) - j(t'_k)| < \delta \qquad (k = 0, 1, \ldots, n-1),$$

i.e. $d_T < \eta \Rightarrow d_T < \delta$. Par conséquent,

$$(\forall\varepsilon > 0)(\exists\eta > 0)(\forall(T', \Xi'))(d_T < \eta \Rightarrow |R^{(g)}_{T,\Xi} - R| = |R^{(f)}_{T,\Xi} - R| < \varepsilon),$$

i.e. $R^{(g)}_{T,\Xi} \to R$ lorsque $d_T \to 0$.

3. Propriétés fondamentales de l'intégrale curviligne

THÉORÈME 1. *Soient L_1 et L_2 deux chemins dans un domaine D plan, tels que l'extrémité de L_1 coïncide avec l'origine de L_2. Si la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ définie sur D est intégrable le long de chacun de ces*

chemins, elle est également intégrable le long de leur somme $L = L_1 + L_2$, *et*

$$\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_{L_1} P\,\mathrm{d}x + Q\,\mathrm{d}y + \int_{L_2} P\,\mathrm{d}x + Q\,\mathrm{d}y.$$

Démonstration. D'après la définition de la somme des chemins, il existe une représentation paramétrique f du chemin L, avec $D_f = [a, c]$, et un point $b \in \,]a, c[$ tels que la restriction f_1 de f au segment $\Delta_1 = [a, b]$ est une représentation paramétrique du chemin L_1 et que la restriction f_2 au segment $\Delta_2 = [b, c]$ est celle du chemin L_2. Soit

$$R_i = \int_{L_i} P\,\mathrm{d}x + Q\,\mathrm{d}y \qquad (i = 1, 2).$$

Fixons un $\varepsilon > 0$. Il existe alors un $\delta_i > 0$ ($i = 1, 2$) tel que pour toute subdivision (T_i, Ξ_i) du segment Δ_i on a

$$d_{T_i} < \delta_i \Rightarrow |R^{(f_i)}_{T_i, \Xi_i} - R_i| < \frac{\varepsilon}{3}.$$

Soit F une application de D dans $\mathbf{R}^2$ de composantes P, Q ; rappelons qu'elles sont supposées continues. Vu que l'application $F \circ f$ du segment $[a, c]$ dans $\mathbf{R}^2$ est alors continue, elle est bornée, i.e. il existe un $C > 0$ tel que

$$\|F(f(t))\| < C \text{ pour tout } t \in [a, c].$$

Etant continue, f est uniformément continue sur $[a, c]$, de sorte qu'il existe un $\delta_3 > 0$ tel que

$$\forall t', t'' \in [a, c], \ |t' - t''| < \delta_3 \Rightarrow \|f(t') - f(t'')\| < \frac{\varepsilon}{9C}.$$

Posons $\delta = \min\{\delta_1, \delta_2, \delta_3\}$ et soit (T, Ξ) une subdivision arbitraire du segment $[a, c]$, avec $d_T < \delta$. Deux cas sont possibles :

1) c est un point de la subdivision T, $c = t_k$. Alors

$$R^{(f)}_{T, \Xi} = R^{(f_1)}_{T_1, \Xi_1} + R^{(f_2)}_{T_2, \Xi_2},$$

où (T_1, Ξ_1) est une subdivision du segment Δ_1 par les points $t_0, \tau_0, t_1, \tau_1, \ldots, t_{k-1}, \tau_{k-1}, t_k$ et (T_2, Ξ_2) est une subdivision du segment Δ_2 par les points $t_k, \tau_k, \ldots, t_{n-1}, \tau_{n-1}, t_n$. Comme $d_{T_1} < \delta \leqslant \delta_1$ et $d_{T_2} < \delta \leqslant \delta_2$, on a

$$|R^{(f)}_{T, \Xi} - (R_1 + R_2)| = |(R^{(f_1)}_{T_1, \Xi_1} - R_1) + (R^{(f_2)}_{T_2, \Xi_2} - R_2)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon;$$

2) il existe un k tel que $t_k < c < t_{k+1}$. Définissons les subdivisions (T_i, Ξ_i) des segments Δ_i de la façon suivante :

$$(T_1, \Xi_1) : t_0, \tau_0, \ldots, \tau_{k-1}, t_k, \tau_k', c;$$
$$(T_2, \Xi_2) : c, \tau_k'', t_{k+1}, \tau_{k+1}, \ldots, \tau_{n-1}, t_n,$$

où τ_k' et τ_k'' sont respectivement des points de $[t_k, c]$ et $[c, t_{k+1}]$. Alors,

$$R_{T,\Xi}^{(f)} = R_{T_1,\Xi_1}^{(f_1)} + R_{T_2,\Xi_2}^{(f_2)} + \langle F(f(\tau_k)), f(t_{k+1}) - f(t_k)\rangle -$$
$$- \langle F(f(\tau_k')), f(c) - f(t_k)\rangle - \langle F(f(\tau_k'')), f(t_{k+1}) - f(c)\rangle.$$

Etant donné que $d_{T_1} < \delta_1$ et $d_{T_2} < \delta_2$, on obtient

$$|R_{T,\Xi}^{(f)} - (R_1 + R_2)| \leqslant |R_{T_1,\Xi_1}^{(f_1)} - R_1| + |R_{T_2,\Xi_2}^{(f_2)} - R_2| +$$
$$+ \|F(f(\tau_k))\|\cdot\|f(t_{k+1}) - f(t_k)\| + \|F(f(\tau_k'))\|\cdot\|f(c) - f(t_k)\| +$$
$$+ \|F(f(\tau_k''))\|\cdot\|f(t_{k+1}) - f(c)\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + 3C\cdot\frac{\varepsilon}{9C} = \varepsilon.$$

Ainsi donc,

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall(T, \Xi))(d_T < \delta \Rightarrow |R_{T,\Xi}^{(f)} - (R_1 + R_2)| < \varepsilon),$$

i.e. $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y$ existe et est égale à

$$R_1 + R_2 = \int_{L_1} P\,\mathrm{d}x + Q\,\mathrm{d}y + \int_{L_2} P\,\mathrm{d}x + Q\,\mathrm{d}y.$$

THÉORÈME 2. *La forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ sur D est intégrable le long du chemin L si et seulement si elle est intégrable le long du chemin inverse $-L$. Ceci étant, on a*

$$\int_{-L} P\,\mathrm{d}x + Q\,\mathrm{d}y = -\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y.$$

Démonstration. Soit f une représentation paramétrique du chemin L et soit (T, Ξ) une subdivision du segment $[a, b] = D_f$ par les points $a = t_0$, $\tau_0, t_1, \ldots, \tau_{n-1}, t_n = b$. La symétrie $\sigma_{a,b}$ transforme (T, Ξ) en une subdivision $(\bar{T}, \bar{\Xi})$ de $[a, b]$ par des points $a = \bar{t}_0, \bar{\tau}_0, \bar{t}_1, \ldots, \bar{\tau}_{n-1}, \bar{t}_n = b$ tels que

$$\sigma_{a,b}(t_k) = \bar{t}_{n-k} \quad (k = 0, 1, \ldots, n),$$
$$\sigma_{a,b}(\tau_k) = \bar{\tau}_{n-k-1} \quad (k = 0, 1, \ldots, n-1)$$

(fig. 66). L'application $\bar{f} = f\circ\sigma_{a,b}$ est, par définition, une représentation paramétrique du chemin $-L$. Soit F une application de D dans $\mathbf{R}^2$ de composantes P, Q et soit $l = n - k - 1$; l parcourt avec k les nombres $0, 1, \ldots, n-1$ (mais dans l'ordre inverse). On a

$$R_{T,\Xi}^{(\bar{f})} = \sum_{k=0}^{n-1} \langle F(\bar{f}(\tau_k)), \bar{f}(t_{k+1}) - \bar{f}(t_k)\rangle =$$
$$= \sum_{k=0}^{n-1} \langle F(f(\bar{\tau}_{n-k-1})), f(\bar{t}_{n-k-1}) - f(\bar{t}_{n-k})\rangle = \sum_{l=0}^{n-1} \langle F(f(\bar{\tau}_l)), f(\bar{t}_l) -$$
$$- f(\bar{t}_{l+1})\rangle = -\sum_{l=0}^{n-1} \langle F(f(\bar{\tau}_l)), f(\bar{t}_{l+1}) - f(\bar{t}_l)\rangle = -R_{\bar{T},\bar{\Xi}}^{(f)}.$$

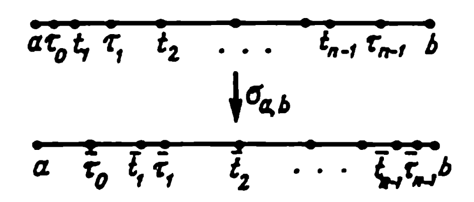


Fig. 66

Si $\int\limits_L P\,dx + Q\,dy$ existe, $R^{(\mathcal{S})}_{\bar T, \bar\Xi} \to \int\limits_L P\,dx + Q\,dy$ lorsque $d_{\bar T} \to 0$. Or $d_T = d_{\bar T}$. Par conséquent, $R^{(\bar{\mathcal{S}})}_{T,\Xi} \to -\int\limits_L P\,dx + Q\,dy$ lorsque $d_T \to 0$, i.e. $\int\limits_{-L} P\,dx + Q\,dy$ existe et est égale à $-\int\limits_L P\,dx + Q\,dy$. Comme $-(-L) = L$, la réciproque est aussi vraie.

Les intégrales curvilignes vérifient l'assertion analogue au critère de convergence de Cauchy.

THÉORÈME 3. *Soient f une représentation paramétrique du chemin L, (T, Ξ) et (T', Ξ') des subdivisions du segment $[a, b] = D_f$ et $R^{(\mathcal{S})}_{T,\Xi}$ et $R^{(\mathcal{S})}_{T',\Xi'}$ les sommes de Riemann respectives de la forme différentielle $P\,dx + Q\,dy$. L'intégrale curviligne $\int\limits_L P\,dx + Q\,dy$ existe si et seulement si*

$$(4)\quad (\forall\varepsilon > 0)(\exists\delta > 0)(\forall(T, \Xi) \text{ et } \forall(T', \Xi'))(d_T < \delta \text{ et } d_{T'} < \delta) \Rightarrow$$
$$\Rightarrow |R^{(\mathcal{S})}_{T,\Xi} - R^{(\mathcal{S})}_{T',\Xi'}| < \varepsilon.$$

Démonstration. 1) La condition (4) est *nécessaire*. En effet, soit $R = \int\limits_L P\,dx + Q\,dy$, de sorte que

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall(T, \Xi))\left(d_T < \delta \Rightarrow |R^{(\mathcal{S})}_{T,\Xi} - R| < \frac{\varepsilon}{2}\right).$$

On a alors

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall(T, \Xi) \text{ et } \forall(T', \Xi'))(d_T < \delta \text{ et } d_{T'} < \delta) \Rightarrow$$
$$\Rightarrow |R^{(\mathcal{S})}_{T,\Xi} - R^{(\mathcal{S})}_{T',\Xi'}| \leqslant |R^{(\mathcal{S})}_{T,\Xi} - R| + |R^{(\mathcal{S})}_{T',\Xi'} - R| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

donc la condition (4) est satisfaite.

2) La condition (4) est *suffisante*. En effet, considérons une suite de subdivisions (T_n, Ξ_n) du segment $[a, b]$ pour laquelle $d_{T_n} \to 0$. Il existe alors pour δ de la condition (4) un N tel que $d_{T_n} < \delta$ pour tout $n > N$,

de sorte que la condition (4) implique :

$$(\forall \varepsilon > 0)(\exists N)(\forall p > N \text{ et } \forall q > N)(|R^{(\mathscr{S})}_{T_p, \Xi_p} - R^{(\mathscr{S})}_{T_q, \Xi_q}| < \varepsilon).$$

Cela signifie que la suite numérique $(R^{(\mathscr{S})}_{T_n, \Xi_n})$ satisfait au critère de Cauchy et, par suite, converge vers un nombre R. Fixons maintenant (T, Ξ) dans la condition (4) et remplaçons (T', Ξ') par (T_n, Ξ_n). En passant à la limite lorsque $n \to +\infty$, on obtient

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall (T, \Xi))(d_T < \delta \Rightarrow |R^{(\mathscr{S})}_{T, \Xi} - R| < \varepsilon),$$

i.e. $R^{(\mathscr{S})}_{T, \Xi} \to R$ lorsque $d_T \to 0$, de sorte que $\int_L P\,dx + Q\,dy$ existe et est égale à R.

Ce théorème est surtout utile lorsqu'on considère la notion de *partie du chemin*. On dit que le chemin L_1 est une partie du chemin L s'il existe une représentation paramétrique f du chemin L et un segment $[a_1, b_1] \subset D_f$ tels que la restriction de f à $[a_1, b_1]$ est une représentation paramétrique du chemin L_1.

THÉORÈME 4. *La forme différentielle $P\,dx + Q\,dy$ intégrable le long du chemin L est intégrable le long de n'importe quelle partie L_1 de L.*

Démonstration. L'assertion du théorème est équivalente à l'affirmation suivante : si la forme différentielle n'est pas intégrable le long d'une partie L_1 du chemin L, elle n'est pas intégrable le long de L non plus. Supposons que f soit une représentation paramétrique du chemin L et que sa restriction f_1 au segment $[a_1, b_1] \subset [a, b] = D_f$ soit une représentation paramétrique du chemin L_1. Si la forme différentielle $P\,dx + Q\,dy$ n'est pas intégrable le long de L_1, il existe en vertu du théorème 3 un $\varepsilon > 0$ tel que pour tout $\delta > 0$ il existe des subdivisions (T_1, Ξ_1) et (T_1', Ξ_1') du segment $[a_1, b_1]$, avec $d_{T_1} < \delta$, $d_{T_1'} < \delta$, pour lesquelles $|R^{(\mathscr{S}_1)}_{T_1, \Xi_1} - R^{(\mathscr{S}_1)}_{T_1', \Xi_1'}| \geqslant \varepsilon$. Prolongeons (T_1, Ξ_1) et (T_1', Ξ_1') par les mêmes points jusqu'aux subdivisions (T, Ξ) et (T', Ξ') du segment $[a, b]$ de telle sorte que les diamètres des subdivisions ne soient pas augmentés. Comme $R^{(\mathscr{S})}_{T, \Xi} - R^{(\mathscr{S})}_{T', \Xi'} = R^{(\mathscr{S}_1)}_{T_1, \Xi_1} - R^{(\mathscr{S}_1)}_{T_1', \Xi_1'}$, on obtient

$$(\exists \varepsilon > 0)(\forall \delta > 0)(\exists (T, \Xi) \text{ et } \exists (T', \Xi'))(d_T < \delta,\ d_{T'} < \delta \text{ et } |R^{(\mathscr{S})}_{T, \Xi} - R^{(\mathscr{S})}_{T', \Xi'}| \geqslant \varepsilon).$$

En vertu du théorème 3 cela signifie que $P\,dx + Q\,dy$ n'est pas intégrable le long de L.

4. Intégrales le long du chemin fermé

Soient L un chemin fermé, f une représentation paramétrique de L, $D_f = [a, b]$ et $c \in]a, b[$. On a alors $L = L_{[a,c]} + L_{[c,b]}$, où $L_{[a,c]}$ est une

partie du chemin L correspondant à la restriction de f à $[a, c]$, et $L_{[c,b]}$ est une partie de L définie de façon analogue. Puisque le chemin L est fermé, $f(a) = f(b)$. C'est pourquoi, il existe le chemin $L_c = L_{[c,b]} + L_{[a,c]}$. Il est en général distinct de L ; par exemple, si $f(c) \neq f(a)$, les chemins L et L_c ont des origines différentes. Cependant, *la forme différentielle* $P\,dx + Q\,dy$ *intégrable le long de* L *est intégrable le long de* L_c, *et*

$$\int_L P\,dx + Q\,dy = \int_{L_c} P\,dx + Q\,dy.$$

En effet, si la forme différentielle $P\,dx + Q\,dy$ est intégrable le long de L, le théorème 4 dit qu'elle est intégrable le long de $L_{[a,c]}$ et le long de $L_{[c,b]}$. Ceci étant, on obtient en vertu du théorème 1 :

$$\int_L P\,dx + Q\,dy = \int_{L_{[a,c]}} P\,dx + Q\,dy + \int_{L_{[c,b]}} P\,dx + Q\,dy.$$

Il résulte toujours du théorème 1 que $P\,dx + Q\,dy$ est intégrable le long de L_c et qu'elle a de plus la même intégrale

$$\int_{L_c} P\,dx + Q\,dy = \int_{L_{[c,b]}} P\,dx + Q\,dy + \int_{L_{[a,c]}} P\,dx + Q\,dy.$$

Ainsi donc, en intégrant le long du chemin fermé, on peut choisir l'origine d'une façon arbitraire (à condition de garder le même sens de parcours).

5. Intégrale curviligne par rapport à l'abscisse et à l'ordonnée

Si $Q \equiv 0$ sur l'ensemble des points du chemin L, on écrit $\int_L P\,dx$ au lieu de $\int_L P\,dx + Q\,dy$ et on dit que cette intégrale curviligne est une *intégrale par rapport à l'abscisse*. De façon analogue, si $P \equiv 0$ sur L, on écrit $\int_L Q\,dy$ au lieu de $\int_L P\,dx + Q\,dy$ et on l'appelle *intégrale par rapport à l'ordonnée.* Soit f une représentation paramétrique du chemin L, avec $D_f = [a, b]$ et les composantes φ et ψ. En utilisant les notations standard, on peut alors écrire :

$$\int_L P\,dx = \lim_{d_T \to 0} \sum_{k=0}^{n-1} P(f(\tau_k))[\varphi(t_{k+1}) - \varphi(t_k)], \tag{5}$$

$$\int_L Q\,dy = \lim_{d_T \to 0} \sum_{k=0}^{n-1} Q(f(\tau_k))[\psi(t_{k+1}) - \psi(t_k)]. \tag{6}$$

Il est évident que

$$\int_L P\,dx + Q\,dy = \int_L P\,dx + \int_L Q\,dy.$$

On voit de la formule (5) que si le chemin L est orthogonal à l'axe x, i.e. φ est constante, on a $\int_L P\,\mathrm{d}x = 0$. De façon analogue, il résulte de la formule (6) que si le chemin L est orthogonal à l'axe y, $\int_L Q\,\mathrm{d}y = 0$. Il résulte aussi de la formule (5) que $P \equiv 1$ est intégrable le long de n'importe quel chemin L et

$$\int_L \mathrm{d}x = \varphi(b) - \varphi(a).$$

De façon analogue, $Q \equiv 1$ est intégrable le long de tout chemin et

$$\int_L \mathrm{d}y = \psi(b) - \psi(a).$$

§ 3. Existence et calcul des intégrales curvilignes

1. Intégrales curvilignes exprimées par une intégrale définie

THÉORÈME 1. *Supposons que la fonction $P(x, y)$ soit continue sur l'ensemble des points du chemin L donné par une représentation paramétrique f, avec $D_f = [a, b]$ et les composantes φ et ψ. Si φ est continûment dérivable, la forme différentielle $P\,\mathrm{d}x$ est intégrable le long de L et*

$$\int_L P(x, y)\,\mathrm{d}x = \int_a^b P(\varphi(t), \psi(t))\varphi'(t)\,\mathrm{d}t. \tag{1}$$

Démonstration. Si $\varphi'(t) \equiv 0$ sur $[a, b]$, le second membre de la formule (1) est nul. Dans ce cas, le premier membre est aussi nul puisque φ est constante. Soit $\varphi'(t) \not\equiv 0$. Comme φ est continue, on a $\int_a^b |\varphi'(t)|\,\mathrm{d}t > 0$. Vu que P est continue sur $f([a, b])$ et f est continue sur $[a, b]$, $P \circ f$ est continue et, par suite, est uniformément continue sur $[a, b]$. Pour cette raison,

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall t \in [a, b] \quad \text{et} \quad \forall \tau \in [a, b]) \left(|t - \tau| < \delta \Rightarrow \right.$$
$$\left. \Rightarrow |P(f(t)) - P(f(\tau))| < \varepsilon \Big/ \int_a^b |\varphi'(t)|\,\mathrm{d}t\right).$$

Si maintenant (T, Ξ) est une subdivision du segment $[a, b]$, on a en notations standard

$$R^{(f)}_{T,\Xi} = \sum_{k=0}^{n-1} P(f(\tau_k))[\varphi(t_{k+1}) - \varphi(t_k)] = \sum_{k=0}^{n-1} P(f(\tau_k)) \int_{t_k}^{t_{k+1}} \varphi'(t)\,\mathrm{d}t =$$
$$= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} P(f(\tau_k))\varphi'(t)\,\mathrm{d}t.$$

D'autre part,

$$\int_a^b P(f(t))\varphi'(t)\mathrm{d}t = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} P(f(t))\varphi'(t)\mathrm{d}t.$$

Si $d_T < \delta$, on a $|t - \tau_k| \leqslant t_{k+1} - t_k < \delta$ pour $t \in [t_k, t_{k+1}]$ (rappelons que $\tau_k \in [t_k, t_{k+1}]$), d'où l'on conclut que

$$\left| \int_a^b P(f(t))\varphi'(t)\mathrm{d}t - R^{(\varphi)}_{T,\Xi} \right| \leqslant \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} |P(f(t)) - P(f(\tau_k))||\varphi'(t)|\mathrm{d}t <$$

$$< \frac{\varepsilon}{\int_a^b |\varphi'(t)|\mathrm{d}t} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} |\varphi'(t)|\mathrm{d}t = \varepsilon.$$

Ainsi donc,

$$\int_a^b P(f(t))\varphi'(t)\mathrm{d}t = \lim_{d_T \to 0} R^{(\varphi)}_{T,\Xi} = \int_L P\,\mathrm{d}x.$$

On démontre de façon analogue le théorème suivant.

THÉORÈME 1'. *Supposons que la fonction $Q(x, y)$ soit continue sur l'ensemble des points du chemin L donné par une représentation paramétrique f, avec $D_f = [a, b]$ et les composantes φ et ψ. Si ψ est continûment dérivable, la forme différentielle $Q\,\mathrm{d}y$ est intégrable le long de L et*

$$(1') \qquad \int_L Q(x, y)\mathrm{d}y = \int_a^b Q(\varphi(t), \psi(t))\psi'(t)\mathrm{d}t.$$

Les théorèmes 1 et 1' impliquent directement le théorème suivant.

THÉORÈME 2. *Si les fonctions $P(x, y)$ et $Q(x, y)$ sont continues sur l'ensemble des points du chemin L de classe C^1, la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ est intégrable le long de L et*

$$(2) \qquad \int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_a^b [P(\varphi(t), \psi(t))\varphi'(t) + Q(\varphi(t), \psi(t))\psi'(t)]\,\mathrm{d}t,$$

où φ et ψ sont les composantes d'une représentation paramétrique continûment différentiable du chemin L définie sur $[a, b]$.

EXEMPLE 1. Soient D un plan de coordonnées à l'origine pointée,

$$P(x, y) = -\frac{y}{x^2 + y^2}, \qquad Q(x, y) = \frac{x}{x^2 + y^2}$$

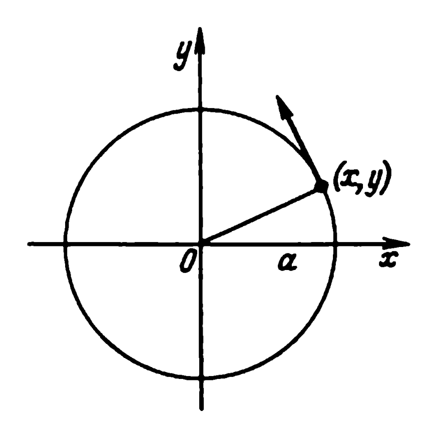


Fig. 67

et L une circonférence de rayon a définie par les équations paramétriques

(3) $$x = a\cos t,\quad y = a\sin t \qquad (0 \leqslant t \leqslant 2\pi).$$

Alors

$$\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_0^{2\pi}\left(-\frac{a\sin t}{a^2}(-a\sin t) + \frac{a\cos t}{a^2}\,a\cos t\right)\mathrm{d}t =$$

$$= \int_0^{2\pi}\mathrm{d}t = 2\pi.$$

On a pu prévoir ce résultat en partant des considérations physiques. En effet, soient P et Q les coordonnées d'un vecteur de longueur $\dfrac{1}{\sqrt{x^2+y^2}}$ perpendiculaire au rayon vecteur du point (x, y). Il s'ensuit (voir n° 2 du § 2) que $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y$ est le travail produit par une force d'intensité constante $1/a$ dirigée le long de la tangente à la circonférence (3), quand le point d'application de cette force parcourt la circonférence (fig. 67) ; il est évident que ce travail est égal au produit de $1/a$ par la longueur $2\pi a$ de la circonférence, i.e. à 2π.

2. Intégrales le long du graphe de la fonction

Soit L le graphe d'une fonction $g(x)$ définie et continue sur un segment non dégénéré $[a, b]$, orienté de gauche à droite, i.e. un chemin défini par les équations paramétriques

$$x = t,\quad y = g(t) \quad (a \leqslant t \leqslant b)$$

(voir n° 3 du § 1). Comme L admet une représentation paramétrique dont la première composante $\varphi(t) = t$ est continûment dérivable (et $\varphi'(t) \equiv 1$), le théorème 1 dit que pour toute fonction $P(x, y)$ continue sur L, la forme

différentielle $P\,dx$ est intégrable le long de L et

$$\int_L P\,dx = \int_a^b P(t,\ g(t))dt = \int_a^b P(x,\ g(x))dx.$$

De façon analogue, si l'on prend l'axe y pour l'axe des abscisses et que L soit le graphe d'une fonction $h(y)$ continue sur un segment non dégénéré $[c,\ d]$, orienté de haut en bas, on a

$$\int_L Q\,dy = \int_c^d Q(h(y),\ y)dy.$$

§ 4. Formule de Green

La formule de Green établit une relation entre les intégrales curvilignes et les intégrales doubles. On ne la démontrera ici que pour quelques cas particuliers.

1. Formule de Green pour les figures simples

1° Soit σ une figure fermée limitée par les droites $x = a$ et $x = b$ et par les graphes des fonctions continues $\varphi_1(x)$ et $\varphi_2(x)$ (fig. 68 ; repère direct). On a vu au n° 11.5.2 que σ est un compact quarrable. Sa frontière est un ensemble des points du chemin fermé

$$L = \smile{AB} + \overline{BC} + \smile{CD} + \overline{DA},$$

où $\smile{AB}$ est le graphe de la fonction φ_1 sur $[a,\ b]$, orienté de gauche à droite, $\overline{BC}$ est le segment d'origine $B(b,\ \varphi_1(b))$ et d'extrémité $C(b,\ \varphi_2(b))$, $\smile{CD}$ est le graphe de la fonction φ_2 sur $[a,\ b]$, orienté de droite à gauche et $\overline{DA}$ est le segment d'origine $D(a,\ \varphi_2(a))$ et d'extrémité $A(a,\ \varphi_1(a))$; les segments $\overline{BC}$ et $\overline{DA}$ peuvent dégénérer en un point.

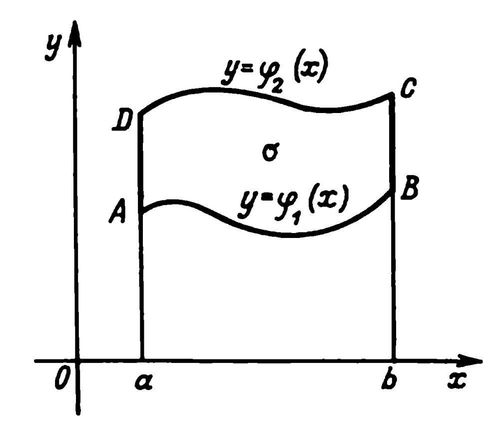


Fig. 68

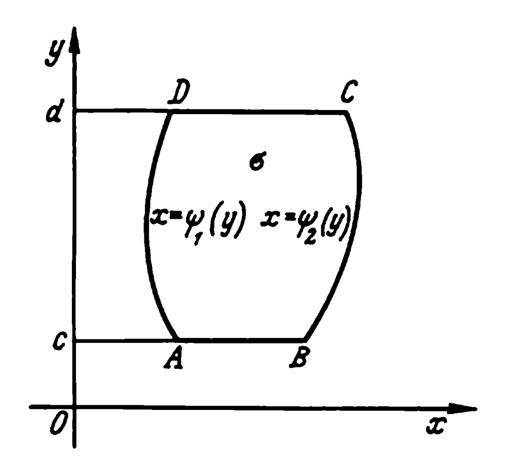


Fig. 69

Soit D un domaine qui contient σ et soit P(x, y) une fonction définie dans D et continue en tous les points du compact σ, ainsi que sa dérivée partielle par rapport à y. Alors la forme différentielle P dx est intégrable le long du chemin L et

(1)
$$\int_L P\,\mathrm{d}x = -\iint_\sigma \frac{\partial P}{\partial y}\,\mathrm{d}x\,\mathrm{d}y.$$

En effet,

$$\iint_\sigma \frac{\partial P}{\partial y}\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}\,\mathrm{d}y\right)\mathrm{d}x =$$

$$= \int_a^b [P(x,\ \varphi_2(x)) - P(x,\ \varphi_1(x))]\,\mathrm{d}x = \int_{\breve{DC}} P\,\mathrm{d}x - \int_{\breve{AB}} P\,\mathrm{d}x =$$

$$= -\int_{\breve{AB}} P\,\mathrm{d}x - \int_{\overline{BC}} P\,\mathrm{d}x - \int_{\breve{CD}} P\,\mathrm{d}x - \int_{\overline{DA}} P\,\mathrm{d}x = -\int_L P\,\mathrm{d}x.$$

2° Soit maintenant σ une figure fermée limitée par les droites $y = c$ et $y = d$ et par les graphes $x = \psi_1(y)$ et $x = \psi_2(y)$ des fonctions continues ψ_1 et ψ_2 (fig. 69) ; σ est un compact quarrable. Sa frontière est un ensemble des points du chemin fermé

$$L = \overline{AB} + \breve{BC} + \overline{CD} + \breve{DA}$$

(la frontière de σ est de nouveau orientée dans le sens direct !).

Soit D un domaine qui contient σ et soit $Q(x, y)$ une fonction définie dans D et continue en tous les points du compact σ, ainsi que sa dérivée partielle par rapport à x. Alors la forme différentielle $Q\,\mathrm{d}y$ est intégrable le long du chemin L et

$$\int_L Q\,\mathrm{d}y = \iint_\sigma \frac{\partial Q}{\partial x}\,\mathrm{d}x\,\mathrm{d}y. \tag{2}$$

En effet,

$$\iint_\sigma \frac{\partial Q}{\partial x}\,\mathrm{d}x\,\mathrm{d}y = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} \frac{\partial Q}{\partial x}\,\mathrm{d}x\right)\mathrm{d}y =$$

$$= \int_c^d [Q(\psi_2(y), y) - Q(\psi_1(y), y)]\,\mathrm{d}y = \int_{\overline{AB}} Q\,\mathrm{d}y + \int_{\overline{BC}} Q\,\mathrm{d}y +$$

$$+ \int_{\overset{\smile}{CD}} Q\,\mathrm{d}y + \int_{\overset{\smile}{DA}} Q\,\mathrm{d}y = \int_L Q\,\mathrm{d}y.$$

REMARQUE. La présence des signes différents dans les formules (1) et (2) s'explique par le fait que le repère et le chemin L sont d'une même orientation directe dans le premier cas, tandis que dans le deuxième, lorsque l'axe y est pris pour l'axe des abscisses et le repère devient rétrograde, ils sont d'orientations contraires.

3° On appelle *figure élémentaire* sur le plan de coordonnées un compact σ qui est en même temps de la forme 1° et de la forme 2° (fig. 70). Par exemple, toute figure convexe est élémentaire (fig. 71).

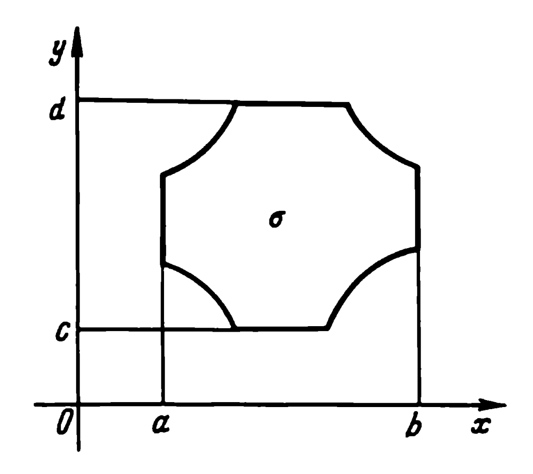


Fig. 70

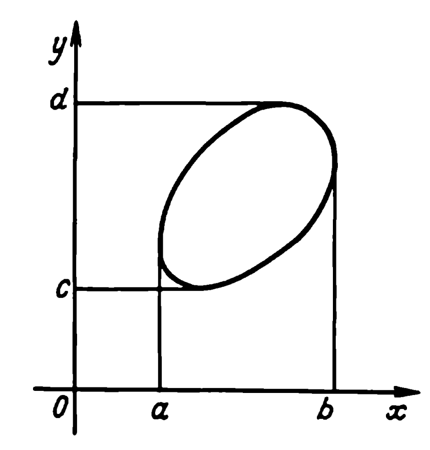


Fig. 71

Supposons que dans un domaine D qui contient une figure élémentaire σ, soient données deux fonctions P(x, y) et Q(x, y) continues et ayant des dérivées partielles continues $\frac{\partial P}{\partial y}$ *et* $\frac{\partial Q}{\partial x}$ *en tous les points de σ. Soit L le chemin qu'on obtient de la frontière de σ en la parcourant dans le sens direct. Alors en vertu des formules* (1) *et* (2), *la forme différentielle* $P\,\mathrm{d}x + Q\,\mathrm{d}y$ *est intégrable le long du chemin L et*

$$\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_\sigma \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y. \tag{3}$$

C'est la *formule de Green*. Son application à d'autres classes de figures σ et sa généralisation au cas multidimensionnel font l'objet de la *théorie des variétés différentiables* et dépassent le cadre de notre cours d'analyse mathématique.

2. Application de la formule de Green au calcul des aires

Soit σ un compact quarrable dont la frontière est l'ensemble des points d'un chemin fermé simple L ; supposons que le repère et le chemin L soient d'une même orientation directe. *Si la formule de Green peut être appliquée à σ et L, on a*

$$\text{aire } \sigma = \frac{1}{2}\int_L x\,\mathrm{d}y - y\,\mathrm{d}x. \tag{4}$$

En effet, prenons pour le domaine D tout le plan et posons $P(x, y) =$ $= -y/2$, $Q(x, y) = x/2$. Alors $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$, d'où

$$\text{aire } \sigma = \iint_\sigma 1 = \iint_\sigma \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y =$$

$$= \int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \frac{1}{2}\int_L x\,\mathrm{d}y - y\,\mathrm{d}x.$$

EXEMPLE 1. Calculons l'aire de la figure fermée σ limitée par l'astroïde

$$x^{2/3} + y^{2/3} = a^{2/3} \quad (a > 0) \tag{5}$$

(fig. 72). C'est une figure élémentaire limitée par les graphes des fonctions continues (donc, une figure quarrable). L'astroïde (5) est un ensemble des points d'un chemin fermé simple L défini par les équations paramétriques

$$x = a\cos^3 t, \qquad y = a\sin^3 t \qquad (0 \leqslant t \leqslant 2\pi).$$

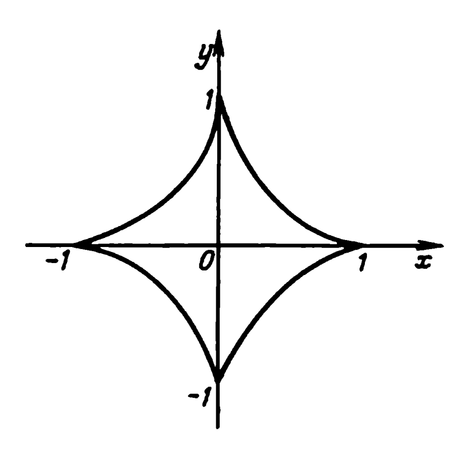


Fig. 72

Ainsi donc, on peut appliquer la formule de Green à σ et L, de sorte que l'aire de σ peut être calculée d'après la formule (4) ; en recourant à la formule (2) du § 3, on obtient

$$\text{aire } \sigma = \frac{1}{2}\int_0^{2\pi}(a\cos^3 t\cdot 3a\sin^2 t\cdot\cos t + a\sin^3 t\cdot 3a\cos^2 t\cdot\sin t)\,dt =$$

$$= \frac{3a^2}{2}\int_0^{2\pi}(\sin^2 t\cos^4 t + \sin^4 t\cos^2 t)\,dt = \frac{3a^2}{2}\int_0^{2\pi}\sin^2 t\cos^2 t\,dt =$$

$$= \frac{3a^2}{8}\int_0^{2\pi}\sin^2 2t\,dt = \frac{3a^2}{8}\int_0^{2\pi}\frac{1-\cos 4t}{2}\,dt = \frac{3}{8}\pi a^2.$$

Ainsi donc, l'aire de l'astroïde est égale à 3/8 de l'aire du disque circonscrit.

Il existe bien sûr une infinité de procédés pour choisir P et Q de telle sorte que $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$. En voici deux avec les formules pour le calcul de l'aire :

$$P = 0, \quad Q = x, \quad \text{aire } \sigma = \int_L x\,dy\,;$$

$$P = -y, \quad Q = 0, \quad \text{aire } \sigma = -\int_L y\,dx.$$

Question. Pourquoi l'intégrale curviligne dans la deuxième formule de l'aire est-elle précédée du signe moins ? Comment s'accorde cela avec ce que l'aire du sous-graphe d'une fonction g positive continue sur le segment $[a, b]$ est égale à $\int_a^b g(x)dx = \int_a^b y\,dx$?

§ 5. Formes différentielles exactes et fermées

1. Intégrales curvilignes dépendant uniquement de l'origine et de l'extrémité du chemin d'intégration

Commençons par un exemple. Soit à calculer le travail produit par une force dans un *champ newtonien*, i.e. dans le champ de forces dues à l'attraction par un point matériel fixé. Ce point sera l'origine O du système de coordonnées cartésiennes rectangulaires xOy (nous n'envisagerons que le cas d'un champ plan). Le domaine de définition D de ce champ est le plan entier pointé en O, et le vecteur défini en tout point de D est une force exercée, d'après la loi de l'attraction universelle, par O sur la masse unité placée en ce point (i.e. une force orientée vers le point O et ayant un module inversement proportionnel au carré de la distance entre ces points).

Soient M un point arbitraire du plan, distinct de O, $\vec{r} = \vec{r}_M$ son rayon vecteur, $r = r_M$ sa longueur et x, y ses coordonnées, de sorte que $r = |\vec{r}| = \sqrt{x^2 + y^2}$. Le point M subit l'action de la force

$$\vec{F} = -k\,\frac{\vec{r}}{r^3},$$

où k est un coefficient strictement positif. En effet, le vecteur $\vec{F}$ est orienté de M vers O, et $|\vec{F}| = k\,\frac{1}{r^2}$, de sorte que le module de la force $\vec{F}$ est inversement proportionnel au carré de la distance de M à O. Par souci de simplicité, choisissons la masse du point O de telle sorte que $k = 1$. Les composantes P, Q du vecteur $\vec{F}$ s'expriment alors par les formules

$$P(x, y) = -\frac{x}{r^3}, \qquad Q = -\frac{y}{r^3}.$$

Soient maintenant A et B des points arbitraires du champ, L un chemin arbitraire de classe C^1 d'origine A et d'extrémité B et R le travail de ce champ le long de L. Alors

$$R = \int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_L -\frac{x\,\mathrm{d}x + y\,\mathrm{d}y}{r^3}.$$

Par hypothèse, le chemin L peut être donné par les équations paramétriques

$$x = \varphi(t), \qquad y = \psi(t) \qquad (a \leqslant t \leqslant b)$$

où φ et ψ sont des fonctions continûment dérivables. Vu que $x^2 + y^2 = r^2$, on a $x\,\mathrm{d}x + y\,\mathrm{d}y = r\,\mathrm{d}r$, d'où

$$-\frac{x\,\mathrm{d}x + y\,\mathrm{d}y}{r^3} = -\frac{\mathrm{d}r}{r^2} = \mathrm{d}\left(\frac{1}{r}\right) = \left(\frac{1}{r}\right)'\mathrm{d}t.$$

Pour cette raison,

$$R = \int_L -\frac{x\,dx + y\,dy}{r^3} = \int_a^b \left(\frac{1}{r}\right)' dt = \frac{1}{r}\Big|_a^b = \frac{1}{r_B} - \frac{1}{r_A},$$

de sorte que R est le même pour tous les chemins L dans D de classe C^1 joignant A et B.

Cette situation se rencontre dans d'autres cas et mérite une généralisation.

On dira que *les intégrales de la forme différentielle* $P\,dx + Q\,dy$ *dans le domaine* D *ne dépendent que de l'origine et de l'extrémité du chemin d'intégration* si pour deux points quelconques A, B de D les intégrales de cette forme le long de tous les chemins de classe C^1 dans D d'origine A et d'extrémité B sont identiques. On a déjà vu que cette propriété est valable pour la forme différentielle

$$-\frac{x\,dx + y\,dy}{(x^2 + y^2)^{3/2}}$$

définie sur le plan de coordonnées pointé à l'origine.

THÉORÈME 1. *Les intégrales de la forme différentielle dans le domaine* D *ne dépendent que de l'origine et de l'extrémité du chemin d'intégration si et seulement si les intégrales de cette forme le long de tous les chemins fermés de classe* C^1 *dans* D *sont nulles.*

Démonstration. Par souci de commodité, désignons la forme différentielle étudiée $P\,dx + Q\,dy$ par la lettre ω.

1) Démontrons que la condition est *nécessaire*. Soit $\int_{L_1} \omega = \int_{L_2} \omega$, quels que soient les chemins L_1, L_2 dans D ayant même origine et même extrémité. Soient L un chemin fermé de classe C^1 dans D, f sa représentation paramétrique continûment différentiable et $D_f = [a, b]$. Si le chemin L contient un seul point, i.e. $f =$ const, on a $\int_L \omega = 0$. Supposons que L contienne plus d'un point. Il existe alors un point $c \in\,]a, b[$ pour lequel $f(c) \neq f(a)$. Soient L_1, L_2 des parties du chemin L, qui correspondent à la restriction de f aux segments $[a, c]$ et $[c, b]$ respectivement. L_1 et $-L_2$ sont des chemins de classe C^1 ayant une origine commune $f(a) = f(b)$ et une extrémité commune $f(c)$. Donc, en vertu de l'hypothèse, $\int_{L_1} \omega = \int_{-L_2} \omega$. Or $L = L_1 + L_2$.

Par conséquent,

$$\int_L \omega = \int_{L_1} \omega + \int_{L_2} \omega = \int_{L_1} \omega - \int_{-L_2} \omega = 0.$$

2) Démontrons que la condition est *suffisante*. Supposons que les intégrales de la forme différentielle ω le long de tous chemins fermés de classe C^1 dans D sont nulles. Soient L_1 et L_2 des chemins de classe C^1 dans D ayant même origine et même extrémité ; $-L_2$ est aussi un chemin de classe C^1 dans D. Comme $L = L_1 + (-L_2)$ est un chemin fermé de classe C^1 dans D, $\int_{L_1}\omega - \int_{L_2}\omega = \int_{L_1}\omega + \int_{-L_2}\omega = \int_{L}\omega = 0$ et, par suite, $\int_{L_1}\omega = \int_{L_2}\omega$.

2. Formes différentielles exactes

Qu'est-ce qui détermine donc que les intégrales de la forme différentielle ne dépendent que des extrémités initiale et terminale du chemin d'intégration ? Dans le cas d'un champ newtonien plan, ce fait résulte de ce que $-\dfrac{x\,\mathrm{d}x + y\,\mathrm{d}y}{r^3} = \mathrm{d}\left(\dfrac{1}{r}\right)$. Il s'avère que la même condition reste valable dans le cas général.

DÉFINITION 1. La forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ sur le domaine D est dite *exacte* s'il existe sur D une fonction différentiable F telle que $P\,\mathrm{d}x + Q\,\mathrm{d}y = \mathrm{d}F$, de sorte que $P = \dfrac{\partial F}{\partial x}$, $Q = \dfrac{\partial F}{\partial y}$. La fonction F est alors appelée *primitive* de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ dans le domaine D.

En particulier, la forme différentielle $-\dfrac{x\,\mathrm{d}x + y\,\mathrm{d}y}{r^3}$ (où $r = \sqrt{x^2 + y^2}$) est exacte dans le plan de coordonnées pointé à l'origine et possède $1/r$ pour sa primitive. (La grandeur $1/r$ est aussi appelée *potentiel* du champ newtonien.)

Rappelons qu'en envisageant les formes différentielles $P\,\mathrm{d}x + Q\,\mathrm{d}y$, on considère que les fonctions $P(x, y)$ et $Q(x, y)$ sont continues, de sorte que d'après le théorème 2 du § 3, la forme $P\,\mathrm{d}x + Q\,\mathrm{d}y$ est intégrable le long de tout chemin de classe C^1.

THÉORÈME 2. *Les intégrales de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ donnée dans le domaine D ne dépendent que des extrémités initiale et terminale du chemin d'intégration de classe C^1 dans D si et seulement si cette forme est exacte dans D.*

Démonstration. 1) La condition est *nécessaire*. En effet, supposons que les intégrales de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ le long de deux lignes polygonales quelconques dans D ayant même origine et même extrémité soient identiques*). Soient encore $M_0(x_0, y_0)$ un point fixe et $M(x, y)$ un point variable dans le domaine D. En vertu du théorème 3.1.9, il existe une ligne polygonale d'origine M_0 et d'extrémité M contenue dans D. Par

*) Leur existence résulte du théorème 2 du § 3 puisque les lignes polygonales sont des chemins de classe C^1.

hypothèse, les intégrales de la forme envisagée le long de toutes les lignes polygonales M_0M sont identiques, i.e. ne dépendent que de M (puisque le point M_0 est fixe). Posons

$$F(x, y) = F(M) = \int_{M_0M} P\,\mathrm{d}x + Q\,\mathrm{d}y$$

et montrons que F est une primitive de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ dans le domaine D. En effet, soient M_1 un point de coordonnées $(x + h, y)$ et $\overline{M_0MM_1}$ une ligne polygonale obtenue par adjonction du segment $\overline{MM_1}$ à M_0M (fig. 73). En vertu de l'additivité de l'intégrale curviligne et du théorème de la moyenne pour l'intégrale définie de la fonction continue, on a

$$\frac{F(x + h, y) - F(x, y)}{h} = \frac{1}{h} \int_{\overline{M_0M_1}} P\,\mathrm{d}x + Q\,\mathrm{d}y =$$

$$= \frac{1}{h} \int_{\overline{MM_1}} P\,\mathrm{d}x = \frac{1}{h} \int_x^{x+h} P(t, y)\mathrm{d}t = P(\xi, y),$$

où ξ est un nombre compris entre x et $x + h$. Par suite,

$$\frac{F(x + h, y) - F(x, y)}{h} \to P(x, y) \quad \text{lorsque} \quad h \to 0,$$

i.e. $\frac{\partial F}{\partial x}$ existe au point (x, y) et est égale à $P(x, y)$. De façon analogue, en joignant le segment $\overline{MM_2}$ à la ligne polygonale M_0M, où M_2 est un point de coordonnées $(x, y + k)$, on obtient que $\frac{\partial F}{\partial y}$ existe au point (x, y) et est

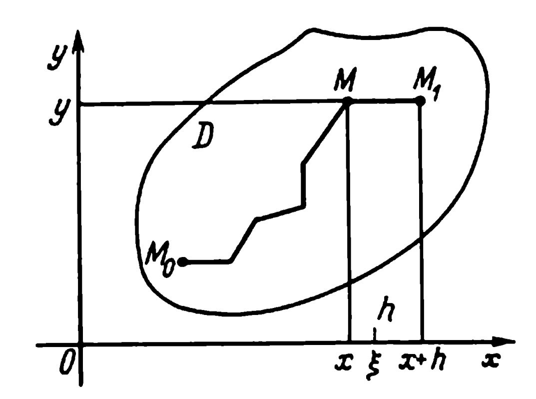


Fig. 73

égale à $Q(x, y)$. Puisque P et Q sont continues, il en résulte que F est différentiable et que $\mathrm{d}F = P\,\mathrm{d}x + Q\,\mathrm{d}y$.

2) La condition est *suffisante*. En effet, soient $P\,\mathrm{d}x + Q\,\mathrm{d}y = \mathrm{d}F$ dans D et L un chemin arbitraire de classe C^1 dans D d'origine A et d'extrémité B. Par définition, il possède une représentation paramétrique continûment différentiable f ; soit $D_f = [a, b]$, de sorte que $f(a) = A$, $f(b) = B$, et soient φ et ψ les composantes de l'application f. Suivant le théorème 2 du § 3,

$$\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_a^b [P(f(t))\varphi'(t) + Q(f(t))\psi'(t)]\,\mathrm{d}t.$$

Or

$$P(f(t))\varphi'(t) + Q(f(t))\psi'(t) = F'_x(f(t))\varphi'(t) + F'_y(f(t))\psi'(t) =$$
$$= F(\varphi(t),\ \psi(t))' = (F \circ f)'(t).$$

Par conséquent,

$$(1) \qquad \int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_a^b (F \circ f)'(t)\,\mathrm{d}t =$$
$$= F(f(b)) - F(f(a)) = F(B) - F(A),$$

de sorte que l'intégrale $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y$ est la même pour tous les chemins L de classe C^1 dans D d'origine A et d'extrémité B.

REMARQUE. La comparaison de deux parties de la démonstration montre que si les intégrales de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ le long de toutes les lignes polygonales ayant même origine et même extrémité sont identiques, les intégrales le long de tous les chemins de classe C^1 ayant même origine et même extrémité sont aussi identiques.

La comparaison des théorèmes 1 et 2 montre que leur corollaire est également juste.

COROLLAIRE. *La forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ donnée dans le domaine D est exacte si et seulement si ses intégrales le long de tous les chemins fermés de classe C^1 dans D sont nulles.*

Il est évident que si F est une primitive de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$, $F + c$ est aussi sa primitive, quelle que soit la constante c.

THÉORÈME 3. *Deux primitives de la forme différentielle exacte ne se distinguent que par une constante.*

Démonstration. Soient F et F_1 des primitives de la forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ dans le domaine D, M_0 un point fixe et M un point variable dans D et soit M_0M une ligne polygonale d'origine M_0 et d'extrémité M contenue dans D (qui existe d'après le théorème 3.1.9). En vertu de la for-

mule (1), on a

$$F_1(M) - F_1(M_0) = \int\limits_{M_0M} P\,dx + Q\,dy = F(M) - F(M_0),$$

d'où $F_1(M) - F(M) = F_1(M_0) - F(M_0) = \text{const.}$

3. Formes différentielles fermées

DÉFINITION 2. La forme différentielle $P\,dx + Q\,dy$ dans le domaine D est dite *fermée* si elle est localement exacte, i.e. tout point de D possède un voisinage dans lequel cette forme est exacte.

THÉORÈME 4. *Soient $P(x, y)$ et $Q(x, y)$ des fonctions continues ayant les dérivées partielles continues $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ dans le domaine D. Pour que la forme différentielle $P\,dx + Q\,dy$ soit fermée, il est nécessaire et suffisant que*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

pour tous les points du domaine D. Si cette condition est accomplie, la forme $P\,dx + Q\,dy$ est exacte dans tout disque ouvert contenu dans D.

Démonstration. 1) La condition est *nécessaire*. En effet, supposons que la forme $P\,dx + Q\,dy$ soit fermée dans D. Tout point $M \in D$ possède alors un voisinage ouvert $U \subset D$ dans lequel cette forme a une primitive F, de sorte que $P = F'_x$, $Q = F'_y$ dans U. Mais par hypothèse, P possède dans D, et donc dans U, une dérivée partielle continue $\frac{\partial P}{\partial y}$. Par conséquent, F admet dans U une dérivée continue F''_{xy}. Il s'ensuit alors du théorème 8.1.1 que $F''_{yx} = (Q'_x) = F''_{xy}$, i.e. $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$.

2) La condition est *suffisante*. En effet, soit $M_0(x_0, y_0)$ un point arbitraire du domaine D. Vu que D est un ensemble ouvert, il existe un disque ouvert $U(M_0\,;\,\varrho)$ de rayon $\varrho > 0$ et de centre M_0 contenu dans D. Ce disque contient avec chacun de ses points $M(x, y)$ un rectangle σ de sommets M_0, $M_1(x, y_0)$, M et $M_2(x_0, y)$ (fig. 74). La forme $P\,dx + Q\,dy$, le rectangle σ et sa frontière $M_0M_1MM_2M_0$ sont justiciables de la formule de Green :

$$\int\limits_{M_0M_1MM_2M_0} P\,dx + Q\,dy = \iint\limits_{\sigma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

Mais par hypothèse, $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. Par conséquent,

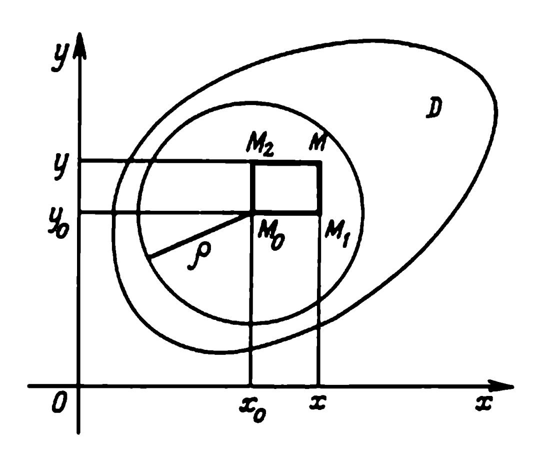


Fig. 74

$$\int_{M_0M_1MM_2M_0} P\,dx + Q\,dy = 0,$$

i.e.

$$\int_{M_0M_2M} P\,dx + Q\,dy = \int_{M_0M_1M} P\,dx + Q\,dy.$$

Soit $F(x, y)$ la valeur commune à deux dernières intégrales. Ainsi donc,

(2) $$F(x, y) = \int_{M_0M_2} Q\,dy + \int_{M_2M} P\,dx = \int_{y_0}^{y} Q(x_0, \eta)d\eta + \int_{x_0}^{x} P(\xi, y)d\xi$$

et

(3) $$F(x, y) = \int_{M_0M_1} P\,dx + \int_{M_1M} Q\,dy = \int_{x_0}^{x} P(\xi, y_0)d\xi + \int_{y_0}^{y} Q(x, \eta)d\eta.$$

Il résulte de la formule (2) que F est différentiable par rapport à x et $F'_x(x, y) = P(x, y)$, et il découle de la formule (3) que F est différentiable par rapport à y et $F'_y(x, y) = Q(x, y)$. Par suite, F est une primitive de la forme $P\,dx + Q\,dy$ dans le disque $U(M_0\,;\,\varrho)$.

Il est évident que toute forme exacte est fermée. Mais la réciproque n'est pas vraie, ce qu'on voit d'après l'exemple cité à la fin du § 3. Dans cet exemple, D est un plan de coordonnées pointé à l'origine O,

$$P(x, y) = -\frac{y}{x^2 + y^2}, \qquad Q(x, y) = \frac{x}{x^2 + y^2}$$

et L est une circonférence de rayon $a > 0$ et de centre O orientée dans le sens direct. Vu que

$$\frac{\partial P}{\partial y}(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial Q}{\partial y}(x, y),$$

la forme $P\,dx + Q\,dy$ est fermée d'après le théorème 4. On a vu que

$$\int_L P\,dx + Q\,dy = 2\pi,$$

de sorte que l'intégrale de la forme $P\,dx + Q\,dy$ le long du chemin fermé L n'est pas nulle et donc, d'après le corollaire aux théorèmes 1 et 2, cette forme n'est pas exacte dans D. Ainsi donc, la forme différentielle $\dfrac{-y\,dx + x\,dy}{x^2 + y^2}$ possède une primitive au voisinage de tout point $M \neq O$, mais on ne peut pas « coller » ces primitives pour en obtenir une seule qui soit définie dans tout le domaine D.

§ 6. Intégrales de la forme différentielle fermée le long des chemins homotopes

1. Chemins homotopes à origine et extrémité fixes

On a vu que les chemins peuvent être paramétrés par des segments non dégénérés $[a, b]$ quelconques. Dans ce paragraphe, *tous les arcs et les chemins seront paramétrés par le même segment* $I = [0, 1]$. Au lieu de $I \times I$ on écrira I^2.

DÉFINITION 1. Soient f_0 et f_1 des arcs paramétrés par I ayant même origine et même extrémité et contenus dans un domaine D. On dit que *l'arc f_0 est homotope à l'arc f_1 dans D* s'il existe une application continue f du carré I^2 dans D vérifiant les conditions suivantes :

(1) $$\forall t \in I,\ f(t, 0) = f_0(t),\ f(t, 1) = f_1(t)\ ;$$

(2) $$\forall \tau \in I,\ f(0, \tau) = f_0(0) = f_1(0),\ f(1, \tau) = f_0(1) = f_1(1).$$

Afin de rendre ces conditions plus claires, posons pour chaque $\tau \in I$

$$f_\tau(t) = f(t, \tau) \qquad (t \in I).$$

Tout f_τ est un arc paramétré. Les conditions (1) signifient qu'il coïncide avec f_0 pour $\tau = 0$ et avec f_1 pour $\tau = 1$. Ainsi donc, f réalise une transformation continue de l'arc f_0 en arc f_1 par l'intermédiaire des arcs f_τ. Les conditions (2) signifient que tous les arcs f_τ, y compris les arcs f_0 et f_1, possèdent même origine et même extrémité (fig. 75).

REMARQUE. Il existe aussi une notion plus générale d'homotopie, sans conditions (2). Mais nous n'en aurons pas besoin.

THÉORÈME 1. *L'homotopie des arcs paramétrés est une relation d'équivalence.*

Démonstration. 1) *L'arc f_0 est homotope à f_0.* L'homotopie est réalisée par la fonction $f(t, \tau) \equiv f_0(t)$.

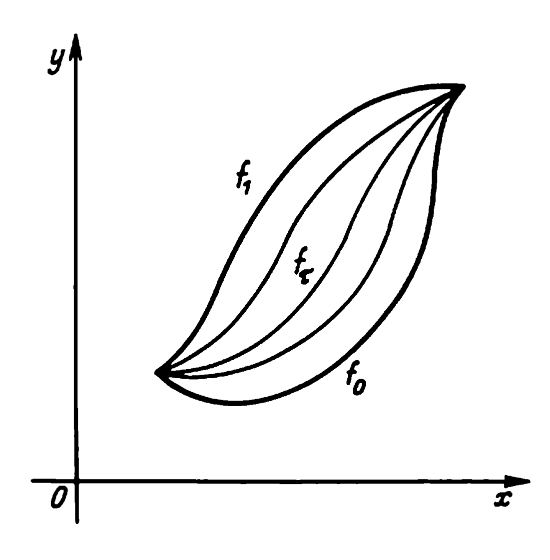


Fig. 75

2) *Si f_0 est homotope à f_1, f_1 est homotope à f_0.* Plus précisément, si f_0 est homotope à f_1 par l'intermédiaire de l'application $f(t, \tau)$, alors f_1 est homotope à f_0 par l'intermédiaire de l'application

$$\bar{f}(t, \tau) = f(t, 1 - \tau).$$

En effet,

$$\bar{f}(t, 0) = f(t, 1) = f_1(t), \bar{f}(t, 1) = f(t, 0) = f_0(t) ;$$
$$\bar{f}(0, \tau) = f(0, 1 - \tau) = f_0(0) = f_1(0),$$
$$\bar{f}(1, \tau) = f(1, 1 - \tau) = f_0(1) = f_1(1).$$

3) *Si f_0 est homotope à f_1 et f_1 est homotope à f_2, f_0 est homotope à f_2.* Soient f et g des applications continues de I^2 dans D telles que f_0 est homotope à f_1 et f_1 est homotope à f_2. Posons pour les points $(t, \tau) \in I^2$

$$h(t, \tau) = \begin{cases} f(t, 2\tau) & \text{si} \quad 0 \leqslant \tau \leqslant 1/2, \\ g(t, 2\tau - 1) & \text{si} \quad 1/2 \leqslant \tau \leqslant 1. \end{cases}$$

Pour $\tau = 1/2$ les deux définitions de $h(t, \tau)$ sont compatibles : $f(t, 1) = f_1(t) = g(t, 0)$ pour tout $t \in I$. Montrons que h est continue. En effet, si $(t_0, \tau_0) \in I \times [0, 1/2[$, il existe en vertu de la continuité de l'application f un voisinage U du point (t_0, τ_0) tel que $U \cap I^2 \subset I \times [0, 1/2[$ et

$$\|h(t, \tau) - h(t_0, \tau_0)\| = \|f(t, 2\tau) - f(t_0, 2\tau_0)\| < \varepsilon$$

pour tout $(t, \tau) \in U \cap I^2$. De façon analogue, si $(t_0, \tau_0) \in I \times]1/2, 1]$, il existe en vertu de la continuité de l'application g un voisinage V du point

(t_0, τ_0) tel que $V \cap I^2 \subset I \times]1/2, 1]$ et

$$\|h(t, \tau) - h(t_0, \tau_0)\| = \|g(t, 2\tau - 1) - g(t_0, 2\tau_0 - 1)\| < \varepsilon$$

pour tout $(t, \tau) \in V \cap I^2$.

Si $t_0 \in I$ et $\tau_0 = 1/2$, il existe, en vertu de la continuité des applications $f(t, 2\tau)$ et $g(t, 2\tau - 1)$ au point (t_0, τ_0), un voisinage W de ce point tel que

$$\|f(t, 2\tau) - f(t_0, 2\tau_0)\| < \varepsilon \text{ pour tout } (t, \tau) \in W \cap (I \times [0, 1/2])$$

et

$$\|g(t, 2\tau - 1) - g(t_0, 2\tau_0 - 1)\| < \varepsilon \text{ pour tout } (t, \tau) \in W \cap (I \times [1/2, 1]),$$

et alors

$$\|h(t, \tau) - h(t_0, \tau_0)\| < \varepsilon \text{ pour tout } (t, \tau) \in W \cap I^2.$$

Enfin, h réalise une homotopie entre f_0 et f_2. En effet,

$$h(t, 0) = f(t, 0) = f_0(t), \qquad h(t, 1) = g(t, 1) = f_2(t),$$
$$h(0, \tau) = f(0, 2\tau) \text{ (ou } = g(0, 2\tau - 1)) = f_0(0) = f_2(0).$$
$$h(1, \tau) = f(1, 2\tau) \text{ (ou } = g(1, 2\tau - 1)) = f_0(1) = f_2(1).$$

En vertu du théorème 1, on peut dire tout simplement que *les arcs* f_0 *et* f_1 *sont homotopes.*

Les arcs f_0, f_1 *définis sur* I *sont homotopes s'ils sont des représentations paramétriques d'un même chemin.* En effet, f_0 et f_1 ont une origine commune (l'origine du chemin L) et une extrémité commune (l'extrémité du chemin L). On a encore $f_1 = f_0 \circ j$ où j est une application strictement croissante du segment I sur lui-même. Si $t, \tau \in I$, il vient $(1 - \tau)t + \tau j(t) \in I$. Posons

$$f(t, \tau) = f_0((1 - \tau)t + \tau j(t)) \qquad ((t, \tau) \in I^2) ;$$

f est une application continue de I^2 dans D, et

$$f(t, 0) = f_0(t), \qquad f(t, 1) = f_0(j(t)) = f_1(t),$$
$$f(0, \tau) = f_0(0) = f_1(0), \qquad f(1, \tau) = f_0(1) = f_1(1),$$

de sorte que f vérifie toutes les conditions pour que f_0 soit homotope à f_1.

Le résultat obtenu permet d'étendre la notion d'homotopie des arcs à celle des chemins.

DÉFINITION 2. On dit que les chemins L_0 et L_1 dans D ayant même origine et même extrémité sont *homotopes* s'il existe des représentations paramétriques f_0 et f_1 de L_0 et L_1 définies sur I qui soient homotopes. (En vertu de cela, toutes leurs représentations définies sur I sont alors homotopes.)

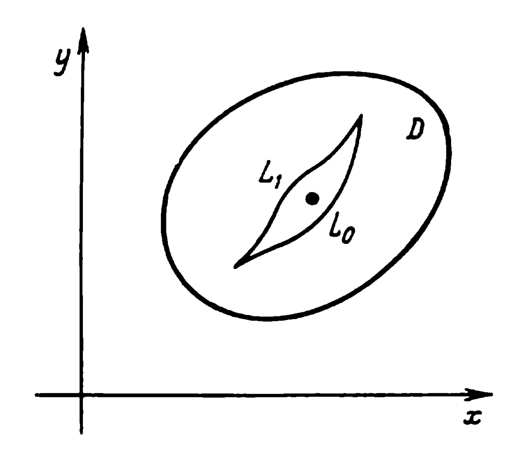


Fig. 76

Dans un domaine convexe, deux chemins quelconques ayant même origine et même extrémité sont homotopes. En effet, l'homotopie de leurs représentations paramétriques f_0 et f_1 définies sur I est réalisée, par exemple, par l'application

$$f(t, \tau) = (1 - \tau)f_0(t) + \tau f_1(t).$$

Il est évident que si l'on exclut un point de ce domaine, l'assertion cesse d'être vraie (fig. 76), ce qui découle du théorème suivant.

2. Intégrales de la forme différentielle fermée le long des chemins homotopes

THÉORÈME 2. *Soient L_0 et L_1 des chemins de classe C^1 dans le domaine D ayant même origine et même extrémité et soient P et Q des fonctions continues sur D ayant des dérivées partielles continues $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$. Si les chemins L_0 et L_1 sont homotopes dans D et la forme différentielle $P\,dx + Q\,dy$ est fermée, on a*

$$\int_{L_0} P\,dx + Q\,dy = \int_{L_1} P\,dx + Q\,dy.$$

Démonstration. Par hypothèse, L_0 et L_1 admettent des représentations paramétriques f_0 et f_1 de classe C^1 définies sur $I = [0, 1]$, et il existe une application continue f du carré I^2 dans D satisfaisant aux conditions (1) et (2). Vu que f est continue et I^2 est un compact, $f(I^2)$ est un compact. Par conséquent, $\varrho = \varrho(f(I^2), \mathrm{Fr}\,D) > 0$ d'après la propriété 3.2.5, 5°, et donc, en vertu du théorème 3.3.4 on a $U(f(C)\,;\,\varrho) \subset D$ pour tout point

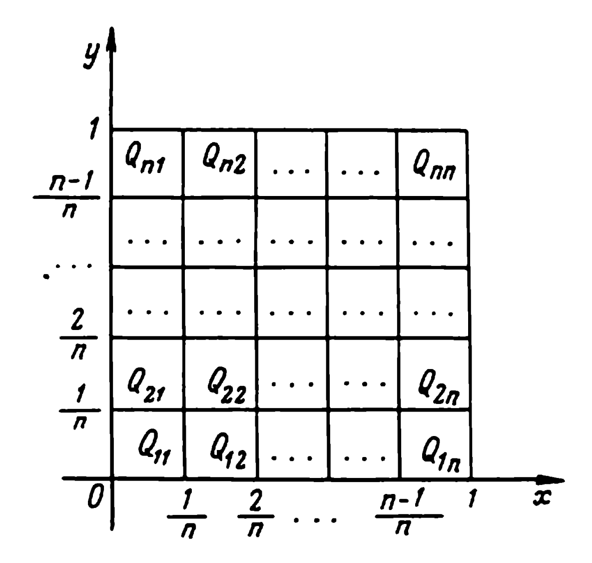


Fig. 77

$C \in I^2$. Le théorème 4 du § 5 dit que la forme $P\,dx + Q\,dy$ possède une primitive sur chaque disque $U(f(C)\,;\,\varrho)$.

Comme f est uniformément continue sur I^2,

$$(\exists\delta > 0)(\forall C \in I^2 \text{ et } \forall C' \in I^2)(d(C, C') < \delta \Rightarrow d(f(C), f(C')) < \varrho).$$

Prenons un entier naturel $n > \frac{1}{\delta\sqrt{2}}$ et subdivisons I^2 en n^2 carrés congruents :

$$\begin{array}{l} Q_{n1},\ Q_{n2},\ \ldots,\ Q_{nn}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ Q_{21},\ Q_{22},\ \ldots,\ Q_{2n}, \\ Q_{11},\ Q_{12},\ \ldots,\ Q_{1n} \end{array}$$

(fig. 77) ; Q_{ij} désigne un carré situé à l'intersection de la i-ième ligne avec la j-ième colonne (à compter à partir des axes de coordonnées). Soit C_{ij} le centre du carré Q_{ij} et soit $U_{ij} = U(f(C_{ij})\,;\,\varrho)$. Vu que diam $Q_{ij} = \frac{1}{n}\sqrt{2} < 2\delta$, on a $d(C_{ij}, C) \leqslant \frac{1}{2}$ diam $Q_{ij} < \delta$ pour tout $C \in Q_{ij}$, d'où $d(f(C_{ij}), f(C)) < \varrho$ et, par conséquent, $f(Q_{ij}) \subset U_{ij}$.

Proposons-nous de construire sur I^2 une fonction g coïncidant sur chaque carré Q_{ij} avec $F_{ij} \circ f|_{Q_{ij}}$, où F_{ij} est une primitive de la forme $P\,dx + Q\,dy$ sur le disque U_{ij} (conformément à ce qu'il a été dit plus haut, cette primitive existe). Notons tout d'abord que $Q_{ij} \cap Q_{i,j+1}$ est un côté commun aux carrés Q_{ij} et $Q_{i,j+1}$ et que

$$f(Q_{ij} \cap Q_{i,j+1}) \subset f(Q_{ij}) \cap f(Q_{i,j+1}) \subset U_{ij} \cap U_{i,j+1},$$

de sorte que les disques ouverts U_{ij} et $U_{i,j+1}$ ont des points communs. Prenons maintenant des primitives arbitraires F_{11} et $\bar{F}_{12}$ de $P\,dx + Q\,dy$ sur U_{11} et U_{12} respectivement. En vertu du théorème 3 du § 5, F_{11} et $\bar{F}_{12}$ se distinguent par une constante sur l'intersection de ces disques : $(\bar{F}_{12} - F_{11})|_{U_{11}\cap U_{12}} = c$. Posons $F_{12} = \bar{F}_{12} - c$. $\bar{F}_{12}$ est aussi une primitive sur U_{12}, et F_{11} et F_{12} coïncident déjà sur $U_{11} \cap U_{12}$. Procédons de la même manière pour déterminer F_{13} sur U_{13} à partir de F_{12}, et ainsi de suite jusqu'à F_{1n} sur U_{1n}. Posons sur chaque carré Q_{1j}

$$g_{1j} = F_{1j} \circ f|_{Q_{1j}}.$$

Puisque $f(Q_{1j} \cap Q_{1,j+1}) \subset U_{1j} \cap U_{1,j+1}$ et que F_{1j} et $F_{1,j+1}$ coïncident sur $U_{1j} \cap U_{1,j+1}$, les fonctions g_{1j} et $g_{1,j+1}$ coïncident aussi sur l'intersection $Q_{1j} \cap Q_{1,j+1}$ de leurs domaines de définition. Il s'ensuit que les fonctions $g_{11}, \ldots, g_{1n}$ sont des restrictions aux carrés respectifs $Q_{11}, \ldots, Q_{1n}$ d'une même fonction définie sur $\bigcup_{j=1}^{n} Q_{1j}$. Maintenant, en montant de chaque carré Q_{1j} le long de la verticale et en partant de F_{1j}, construisons successivement les fonctions $g_{2j} = F_{2j} \circ f|_{Q_{2j}}, \ldots, g_{nj} = F_{nj} \circ f|_{Q_{nj}}$, où F_{ij} est une primitive de $P\,dx + Q\,dy$ sur U_{ij}. Ceci étant, F_{ij} et $F_{i+1,j}$ coïncident sur $U_{ij} \cap U_{i+1,j}$, de sorte que les fonctions g_{ij} et $g_{i+1,j}$ coïncident aussi sur l'intersection $Q_{ij} \cap Q_{i+1,j}$ de leurs domaines de définition. Considérons ensuite, sur $Q_j = \bigcup_{i=1}^{n} Q_{ij}$, des fonctions g_j telles que $g_j|_{Q_{ij}} = g_{ij}$ et montrons que g_j et g_{j+1} coïncident sur l'intersection $Q_j \cap Q_{j+1}$ de leurs domaines de définition. En effet,

$$Q_j \cap Q_{j+1} = \bigcup_{i=1}^{n} (Q_{ij} \cap Q_{i,j+1})\ ;$$

les fonctions g_j et g_{j+1} coïncident sur le segment $Q_{1j} \cap Q_{1,j+1}$ d'après la construction initiale ; par conséquent, elles coïncident également à l'extrémité supérieure de ce segment, qui est en même temps l'extrémité inférieure du segment $Q_{2j} \cap Q_{2,j+1}$; or les primitives F_{2j} et $F_{2,j+1}$ se distinguent sur $f(Q_{2j} \cap Q_{2,j+1})$ par une constante, ce qui entraîne la même chose pour les fonctions g_{2j} et $g_{2,j+1}$ sur $Q_{2j} \cap Q_{2,j+1}$; il s'ensuit que ces fonctions coïncident sur tout ce segment car elles coïncident en son extrémité inférieure ; ainsi donc, les fonctions g_j et g_{j+1} coïncident, elles aussi, sur le segment $Q_{2j} \cap Q_{2,j+1}$. En raisonnant de la sorte, on obtient que g_j et g_{j+1} coïncident sur tout le segment $Q_j \cap Q_{j+1}$.

On a construit ainsi sur I^2 une fonction g telle que

$$g|_{Q_{ij}} = g_{ij} = F_{ij} \circ f|_{Q_{ij}} \quad \text{pour tout } Q_{ij}.$$

La fonction g est constante sur les côtés latéraux du carré I^2. En effet, le côté gauche est composé des points $(0, \tau)$ $(0 \leqslant \tau \leqslant 1)$; si $\frac{i-1}{n} \leqslant i \leqslant \frac{i}{n}$, on a $(0, \tau) \in Q_{i1}$ et donc

$$g(0, \tau) = F_{i1}(f(0, \tau)) = F_{i1}(f_0(0)),$$

i.e. $g(0, \tau)$ est constante sur chaque segment $\{0\} \times \left[\frac{i-1}{n}, \frac{i}{n}\right]$, et comme les segments voisins ont des points communs, $g(0, \tau)$ est constante sur tout le segment $\{0\} \times I$, i.e. sur le côté gauche du carré I^2. De façon analogue, en considérant la fonction $g(1, \tau)$, on s'assure qu'elle est constante sur le côté droit.

Soient maintenant $A_0, A_1, \ldots, A_{n-1}$ des points de subdivision du côté inférieur, i.e. $A_j = (j/n, 0)$. Comme $f_0(t) = f(t, 0)$ est une représentation paramétrique continûment différentiable du chemin L_0, la forme $P\,dx + Q\,dy$ est intégrable le long de L_0 et

$$\int_{L_0} P\,dx + Q\,dy = \sum_{j=1}^{n} \int_{L_0|\overline{A_{j-1}A_j}} P\,dx + Q\,dy.$$

Mais comme $A_{j-1}A_j \subset Q_{1j}$, on a $f(A_{j-1}A_j) \subset U_{1j}$, et puisque $P\,dx + Q\,dy = dF_{1j}$ sur U_{1j}, on peut écrire

$$\int_{L_0} P\,dx + Q\,dy = \sum_{j=1}^{n} \int_{L_0|\overline{A_{j-1}A_j}} dF_{1j} = \sum_{j=1}^{n} [F_{1j}(f(A_j)) - F_{1j}(f(A_{j-1}))] =$$

$$= \sum_{j=1}^{n} \left[g\left(\frac{j}{n}, 0\right) - g\left(\frac{j-1}{n}, 0\right)\right] = g(1, 0) - g(0, 0).$$

En utilisant les points $(i/h, 1)$ de la subdivision du côté supérieur du carré I^2 et la représentation paramétrique $f_1(t) = f(\tau, 1)$ du chemin L, on obtient de même

$$\int_{L_1} P\,dx + Q\,dy = g(1, 1) - g(0, 1).$$

Or, $g(0, 0) = g(0, 1)$ et $g(1, 0) = g(1, 1)$. Par conséquent,

$$\int_{L_0} P\,dx + Q\,dy = \int_{L_1} P\,dx + Q\,dy.$$

3. Domaines simplement connexes

DÉFINITION 3. On dit que le domaine plan D est *simplement connexe* si tous les chemins dans D ayant même origine et même extrémité sont homotopes deux à deux.

Ainsi donc, il a été montré à la fin du n° 2 que *tout domaine convexe est simplement connexe.*

THÉORÈME 3. *Soit Φ une application régulière d'un domaine D du plan de coordonnées Π dans le plan de coordonnées Π_1 et soit $D_1 = \Phi(D)$. Si D est simplement connexe, D_1 l'est aussi.*

Démonstration. Conformément à la remarque au théorème 12.1.1, D_1 est aussi un domaine. Soient L_0 et L_1 deux chemins quelconques dans D_1 ayant même origine et même extrémité et soient f_0 et f_1 des représentations paramétriques de ces chemins, définies sur le segment $I = [0, 1]$. Posons $\bar{f}_i = \Phi^{-1} \circ f_i$ $(i = 1, 2)$; $\bar{f}_0$ et $\bar{f}_1$ sont des arcs dans D ayant même origine et même extrémité. Donc, si le domaine D est simplement connexe, ces arcs sont homotopes, i.e. il existe une application continue $\bar{f}$ du carré I^2 dans D vérifiant les conditions

$$\bar{f}(t, 0) \equiv \bar{f}_0(t), \qquad \bar{f}(t, 1) \equiv \bar{f}_1(t),$$
$$\bar{f}(0, \tau) \equiv \bar{f}_0(0) = \bar{f}_1(0), \qquad \bar{f}(1, \tau) \equiv \bar{f}_0(1) = \bar{f}_1(1).$$

Il s'ensuit que $f = \Phi \circ \bar{f}$ est une application continue de I^2 dans D_1 satisfaisant, évidemment, aux conditions

$$f(t, 0) \equiv f_0(t), \qquad f(t, 1) \equiv f_1(t),$$
$$f(0, \tau) \equiv f_0(0) = f_1(0), \qquad f(1, \tau) \equiv f_0(1) = f_1(1),$$

i.e. réalisant une homotopie des chemins L_0 et L_1 dans D_1. Ainsi donc, la connexité simple de D entraîne celle de D_1.

REMARQUE. La régularité de Φ a été utilisée seulement pour établir que D_1 était un domaine tout comme D. Or le théorème de Brouwer dit que l'image d'un domaine de $\mathbf{R}^n$ par toute application homéomorphe est un domaine. C'est pourquoi, le théorème 3 reste en vigueur si au lieu de la régularité on exige que Φ soit homéomorphe. En particulier, *toute image homéomorphe d'un domaine plan convexe est un domaine simplement connexe* (en général, non convexe).

EXEMPLE 1. Soient K une couronne ouverte $0 \leqslant r_1^2 < x^2 + y^2 < r_2^2$ dans le plan de coordonnées Π_1, D_1 un domaine obtenu de K si on en exclut l'intervalle $]r_1, r_2[$ de l'axe x (fig. 78) et soit D un rectangle ouvert $]r_1, r_2[\times]0, 2\pi[$ du plan Π muni des coordonnées cartésiennes rectangulaires r, φ. Alors $D_1 = \Phi(D)$, où Φ est une application de D dans Π_1 définie par

$$(r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

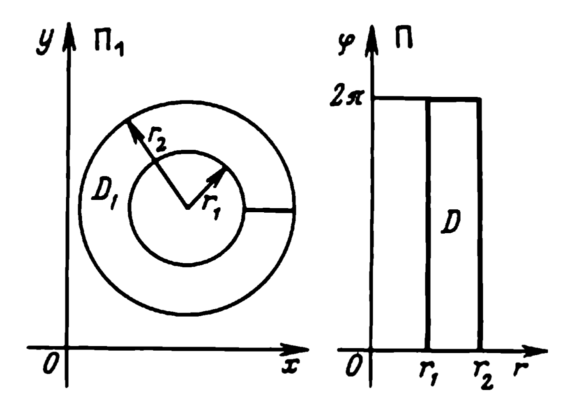


Fig. 78

On sait que Φ est régulière. Vu que le domaine D est simplement connexe parce que convexe, le théorème 3 dit que le domaine D_1 est aussi simplement connexe. Cependant, la couronne K n'est pas simplement connexe. En effet, soit $a \in]r_1, r_2[$, et soient L_0 et L_1 les demi-circonférences supérieure et inférieure :

$$x^2 + y^2 = a^2, \quad y \geqslant 0 \qquad \text{et} \qquad x^2 + y^2 = a^2, \quad y \leqslant 0,$$

orientées de droite à gauche et $L = L_0 + (-L_1)$ (i.e. L est une circonférence $x^2 + y^2 = a^2$ d'origine $(a, 0)$, orientée dans le sens direct). Si K était simplement connexe, les chemins L_0 et L_1 seraient homotopes et pour toute forme différentielle $P\,\mathrm{d}x + Q\,\mathrm{d}y$ sur K satisfaisant aux hypothèses du théorème 2, on aurait

$$\int_{L_0} P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_{L_1} P\,\mathrm{d}x + Q\,\mathrm{d}y,$$

i.e. $\displaystyle\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$. Cependant, la forme $\dfrac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2 + y^2}$ étudiée dans l'exemple 1 du § 3 satisfait aux hypothèses du théorème 2, mais son intégrale n'est pas nulle : $\displaystyle\int_L \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2 + y^2} \neq 0 \ (= 2\pi)$.

EXERCICE 1. Montrer que le domaine pointé n'est pas simplement connexe.

THÉORÈME 4. *Soient P et Q des fonctions continues sur le domaine D et possédant des dérivées partielles continues $\dfrac{\partial P}{\partial y}$ et $\dfrac{\partial Q}{\partial x}$. Si D est simplement connexe, l'égalité $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$ est non seulement nécessaire (en vertu*

du théorème 4 *du* § 5) *mais aussi suffisante pour que la forme différentielle* $P\,\mathrm{d}x + Q\,\mathrm{d}y$ *soit exacte sur* D.

Démonstration. En vertu du corollaire aux théorèmes 1 et 2 du § 5, il suffit de montrer que $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$ pour tout chemin fermé L de classe C^1 dans D. Supposons que L soit un tel chemin. S'il est composé d'un seul point, l'assertion est évidente. S'il contient plus d'un point, on choisit un point $B \in L$ distinct de l'origine A et on le représente sous forme de $L = L_0 + (-L_1)$ où L_0 et L_1 sont des chemins de classe C^1 d'origine A et d'extrémité B. Puisque D est simplement connexe, L_0 et L_1 sont homotopes. La forme $P\,\mathrm{d}x + Q\,\mathrm{d}y$ étant fermée d'après le théorème 4 du § 5, on a, en vertu du théorème 2,

$$\int_{L_0} P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_{L_1} P\,\mathrm{d}x + Q\,\mathrm{d}y$$

d'où $\int_L P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$.

INDEX

